

(19)



Europäisches Patentamt

Europäische Patentbehörde

Office européen des brevets

(11)



EP 0 773 290 A2

(12)

EUROPEAN PATENT APPLICATION

(43) Date of publication:

14.05.1997 Bulletin 1997/20

(51) Int. Cl.⁶: C12N 15/12, C07K 14/705

(21) Application number: 961116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States:

AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

• Arai, Koichi

Urawa-shi, Saitama (JP)

• Yamazaki, Hiroyuki

Higashimurayama-shi, Tokyo (JP)

(30) Priority: 09.10.1995 JP 261440/95
24.04.1996 JP 102451/96

(74) Representative: Wächtershäuser, Günter, Prof.

Dr.

Patentanwalt,

Tal 29

80331 München (DE)

(71) Applicant: KOWA COMPANY, LTD.

Naka-ku Nagoya-shi Aichi-ken (JP)

(72) Inventors:

- Saito, Yasushi
Chiba-shi, Chiba (JP)
- Iwasaki, Akio
Tsuchiura-shi, Ibaraki (JP)

Remarks

The applicant has subsequently filed a sequence listing and declared, that it includes no new matter.

(54) Novel LDL receptor analog protein and the gene coding therefor

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No.5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

EP 0 773 290 A2

Description**Background of the Invention:**5 1) **Field of the Invention**

The present invention relates to a novel LDL receptor analog protein having a structure similar to that of LDL receptors that are responsible for the homeostasis mechanism of intracellular cholesterol and extensively participates in serum lipid metabolism, which is a critical factor that triggers the onset of arteriosclerosis. The invention also relates to 10 the gene coding for the protein.

2) **Description of the Related Art**

Abnormality in serum lipid metabolism is one of the most critical risk factors in the onset and progress of arteriosclerosis. Serum lipids, together with apolipoproteins, are transformed into lipoproteins primarily in the liver, secreted 15 therefrom, transported by blood, and taken up by a variety of tissue cells.

Uptake of lipoproteins into cells occurs primarily by the mediation of receptors of respective lipoproteins. It is known that low density lipoproteins (LDL), which are taken into cells by specific membrane receptors, called LDL receptors, are metabolized within the cells and utilized as cell membrane components or similar substances. Detailed analysis of 20 familial hypercholesterolemia, which is a genetic disease accompanied by notable hypercholesterolemia due to abnormality of LDL receptors, has clarified details of the mechanism of homeostasis achieved by LDL receptors with respect to intracellular cholesterol.

It has been suggested that living bodies have not only LDL receptors but also cell membrane receptors that recognize other lipoproteins. From analyses of WHHL rabbits, which are model animals lacking LDL receptors, it was found 25 that receptors which takes principally apo-E-containing lipoproteins as ligands (remnant receptors) are present in the liver. It is also predicted that there may be HDL receptors whose ligands are high density lipoprotein (HDL). However, to date, details of the structures and functions of these receptors have not yet been elucidated. It has also been known that foaming of macrophages plays an active role in the formation of atherosclerosis, is deeply participated. Macrophages foam by taking up modified LDL—not normal LDL—which have undergone oxidation, acetylation, or glycation. 30 There have recently been discovered receptors to modified LDL which are called scavenger receptors. The scavenger receptors have been identified to be membrane receptors that have a structure completely different from that of LDL receptors.

Recent research using molecular biological techniques has identified the genes of LRP (LDL receptor-associated protein), gp 330, and VLDL receptors. The receptors have been found to have structures very similar to those of LDL 35 receptors. From analyses of these receptors, it is believed that a plurality of lipoprotein receptors are present in living bodies, and that they are closely related to lipid metabolism. LDL receptors studied in detail by Brown and Goldstein [Brown, M.S. and Goldstein, J.L. (1986) *Science* 232, 34-47] are known to play an important role in the homeostasis of lipoprotein metabolism *in vivo*, recognizing apo-B-100 and apo-E and taking primarily LDL as their ligands. Also, LRP, which is a macroprotein, has been found to primarily recognize apo-E and to take β -VLDL or chylomicron remnant as 40 a ligand. Moreover, it has been recently reported that LRP takes an α_2 -macroglobulin/protease complex or a plasminogen activator/plasminogen activator inhibitor-1 complex as a ligand, and that LRP is a protein identical to the α_2 -macroglobulin receptor. When these findings are taken together, LRP is considered to have a wide variety of functions in living bodies [Herz, J., Hamann, U., Rogne, S., Myklebost, O., Gausepohl, H. and Stanley, K.K. (1989) *EMBO J.* 7(13), 4119-4127. Brown, M.S., Herz, J., Kowal, R.C. and Goldstein, J.L. (1991) *Current Opinion in Lipidology* 2, 65-72. Herz, 45 J. (1993) *Current Opinion in Lipidology* 4, 107-113]. The gp 330, which was first identified as an antigen inducing rat Heymann nephritis, has been reported to have a ligand-binding capacity similar to that possessed by CRP α_2 -macroglobulin receptor [Raychowdhury, R., Niles, J.L., McCluskey, R.T. and Smith, J.A. (1989) *Science* 244, 1163-1165. Pietromonaco, S., Kerjaschki, D., Binder, S., Ullrich, R. and Farquhar, G. (1990) *Proc. Natl. Acad. Sci. U.S.A.* 87, 1811-50 1815]. In addition, recently discovered VLDL receptors, which are found to take VLDL as a ligand, are considered to have new functions including fatty acid metabolism, because they are predominantly found in tissues of the heart and muscles though they are rarely found in the liver [Takahashi, S., Kawarabayashi, Y., Nakai, T., Sakai, J. and Yamamoto, T. (1992) *Proc. Natl. Acad. Sci. USA* 89, 9252-9256].

Functions of these newly found receptors as lipoprotein receptors have been gradually elucidated through detailed 55 *in vitro* analyses. However, significance of respective receptors in living bodies has mostly been left unknown. In addition, relations to remnant receptors, HDL receptors, etc., which have conventionally been identified or suggested by biochemical techniques, remain unknown. Presently, it is considered that these newly found receptors are products of genes different from those of the latter receptors. Thus, more lipoprotein receptors than originally guessed have become considered to participate in lipoprotein uptake into cells while interacting with each other to thereby function to maintain homeostasis of lipid metabolism in living bodies. However, from structural analyses of the genes of the afore-

mentioned newly-identified receptors, it is predicted that the genes of these receptors that take lipoproteins as ligands are developed from the same gene from which LDL receptors was developed, and thus they are within the same genetic family. This suggests that lipoprotein receptors that have conventionally been proposed may have structures similar to those of LDL receptors.

5 Accordingly, an object of the present invention is to provide the gene of a novel receptor in the LDL receptor family, as well as a protein coded by the gene.

The present inventors conducted careful studies so as to attain the above object, and found that by using part of rabbit LDL receptor cDNA as a probe there can be obtained a DNA fragment coding for a peptide having a structure similar to that of LDL receptors. Moreover, when using part of the obtained cDNA as a probe, a cDNA fragment having a sequence similar to that of the cDNA can be obtained from the human tissue cDNA library. The present invention was 10 accomplished based on these findings.

Summary of the Invention

15 The present invention provides DNA having a nucleotide sequence shown by Sequence ID No. 1 or No. 5, an LDL receptor analog protein having an amino acid sequence coded by the DNA, a recombinant vector comprising the DNA and a replicable vector; transformant cells which harbor the recombinant vector; and a method for the production of the LDL receptor analog protein.

20 Description of Preferred Embodiment

The cDNA of the present invention may be prepared, for example, by the following process.

Briefly, the process includes the following steps. (1) Through the use of rabbit LDL receptor cDNA as a probe, positive clones are screened out of a rabbit liver cDNA library. (2) Recombinant DNA is prepared using the separated positive clones, and a cDNA fragment is cut out of the resultant recombinant DNA through a treatment using a restriction enzyme. The cDNA fragment is integrated into a plasmid vector. (3) Host cells are transformed using the obtained cDNA recombinant vector to thereby obtain transformant cells of the present invention. The obtained transformant cells are incubated so as to obtain a recombinant vector containing a DNA fragment of the present invention. The nucleotide sequence of the DNA fragment of the present invention contained in the resultant recombinant vector is determined. (4) 25 In tissue of a living body, there is detected expression of mRNA indicated by the nucleotide sequence of the cDNA of the present invention by using RNA blot hybridization method. (5) Through use of a rabbit cDNA fragment as a probe, positive clones are screened out of a human tissue cDNA library, and the nucleotide sequence of the clones is determined. (6) A recombinant vector for expression is prepared using the cDNA of the present invention. Through use of the thus-obtained vector, host cells are transformed to thereby obtain the transformants of the present invention. (7) Ligands that are bound to protein expressed by the obtained transformants are detected by ligand blotting.

35 Each of the above-described steps will next be described.

(1) Screening for positive clones from a rabbit liver cDNA library:

40 A cDNA library may be prepared by the use of mRNA obtained from rabbit liver, reverse transcriptase, and a suitable vector, e.g., commercially available λ gt10 vector.

A cDNA library thus prepared using λ gt10 as a vector is subjected to a screening for positive clones by the application of a DNA hybridization method employing a cDNA probe, to thereby separate positive clones [Sambrook, J., Fritsch, E F and Maniatis, T. (1989) In: Molecular Cloning. A Laboratory Manual, pp 9.47-9.58, Cold Spring Harbor Laboratory Press].

45 An exemplary cDNA which may be used as a probe is rabbit LDL receptor cDNA. Positive clones may be detected by autoradiography employing a DNA probe labelled with a radioisotope (32 P).

(2) Preparation of a cDNA recombinant vector:

50 Recombinant vector λ gt10 phage DNA is extracted from the isolated positive clones and purified. The resultant purified recombinant vector λ gt10 phage DNA is digested with a restriction enzyme EcoRI, to thereby separate a cDNA fragment from the vector DNA. The obtained cDNA fragment is integrated with a plasmid vector for cloning that has been similarly digested with EcoRI, thereby obtaining a recombinant plasmid vector. An exemplary plasmid vector 55 which may be used is pBluescript II.

(3) Recombinant vector, transformation of host cells using the recombinant vector, and preparation of DNA:

The obtained cDNA recombinant vector is introduced into a variety of host cells that are capable of utilizing the

genetic marker possessed by the recombinant vector, to thereby transform the host cells. Host cells are not particularly limited, with *E. coli* being preferred. For example, a variety of variants of the *E. coli* K12 strain, e.g., HB-101, may be used. In order to introduce the recombinant vector into host cells, a competent cell method may be used in combination with a treatment with calcium.

5 The thus-obtained transformant cells are cultured in a selective medium in accordance with the genetic marker of the vector. The recombinant vector of the present invention is collected from the cultured cells. The DNA nucleotide sequence of the cDNA contained in the obtained recombinant vector can be determined through use of a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467].

10 (4) RNA blot hybridization:

The expression in tissue of mRNA, indicated by the nucleotide sequence of the cDNA of the present invention, is detected using RNA blot hybridization.

First, mRNA is prepared using rabbit tissue. Commercially available oligo(dT)cellulose column may be used for the preparation, in order to prepare mRNA from human tissue, there may be used a commercially available nylon membrane on which tissue poly(A)⁺RNA from a variety of sources is present.

An exemplary probe is the rabbit cDNA obtained in the above-described step (3). mRNA may be detected by autoradiography employing a DNA probe labelled with a radioisotope (³²P)

20 (5) Screening of human tissue cDNA library for positive clones, and determination of nucleotide sequence:

An exemplary human tissue cDNA library which may be used is a commercially available human brain cDNA library.

25 Screening and nucleotide sequencing of the human brain cDNA library may be performed using a fragment of rabbit cDNA of the present invention as a probe in a manner similar to that used for the aforementioned rabbit liver cDNA library.

(6) Preparation of a recombinant vector for expression and transformation of host cells using the recombinant vector for expression:

30 In order to prepare an LDL receptor analog protein through use of cDNA of the present invention, the obtained cDNA and a vector for expression are first bonded to each other to thereby create a recombinant vector for expression. Vectors for expression which may be used for bonding are not particularly limited. For example, pBK-CMV may be used.

35 Host cells are transformed using the thus-obtained recombinant vector for expression, to thereby obtain a transformant cell of the present invention. The obtained transformant cell is cultured so as to obtain cells that are capable of expressing the protein of the invention. Host cells are not particularly limited. For example, CHO cells may be used. In order to introduce the recombinant vector for expression into host cells, a calcium phosphate method may be used.

The thus-prepared transformant cells are incubated in a selective medium in accordance with the genetic marker of the vector, so as to express the LDL receptor analog protein of the present invention.

40 (7) Ligand analysis of the protein by ligand blotting:

After the resultant transformant cells are incubated, the expressed LDL receptor analog protein is solubilized using a solubilizer, e.g., Triton X-100, to thereby obtain a membrane protein fraction. The fraction is separated using SDS-PAGE, and transferred onto, for example, a nitrocellulose membrane. Using a radio-labelled (¹²⁵I) lipoprotein as a probe, the analog protein can be detected by autoradiography. Exemplary lipoproteins which may be used include β -VLDL and LDL.

45 Examples:

50 The present invention will next be described in detail by way of example, which should not be construed as limiting the invention.

Example 1:

55 Preparation of a rabbit liver cDNA library:

From tissue of the liver of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to

thereby obtain purified poly(A)⁺RNA.

5 cDNA was synthesized in accordance with a method of Gubler and Hoffman [Gubler, U and Hoffman, B.J. (1983) Gene 25, 263]. Briefly, cDNA was synthesized employing rabbit liver poly(A)⁺RNA (as a template), a random primer, and moloney murine leukemia virus reverse transcriptase. The synthesized cDNA was transferred into double-stranded DNA using DNA polymerase I, and then subjected to an EcoRI methylase treatment. By the use of T4 DNA polymerase, the DNA was blunt-ended. The blunt-ended DNA was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) using a T4 DNA ligase, and the resultant ligated product was subjected to an additional digestion with EcoRI. cDNA fragments having a size not less than 1 kb were selected by agarose gel electrophoresis, and integrated into the EcoRI-digested site of λgt10 phage DNA using a T4 DNA ligase. The phage DNA was packaged *in vitro*, to thereby establish a rabbit liver cDNA library.

10 Example 2:

15 Cloning of cDNA of receptors in the rabbit LDL receptor family:

20 The cDNA library (1,000,000 plaques) prepared in Example 1 was subjected to screening using a plaque hybridization method and employing as a probe a segment of the cDNA obtained from a ligand binding region, the functional region, of the rabbit LDL receptor. Hybridization was performed at 42°C using 5 x SSC, 30% formamide, 1% SDS, 5 x Denhardt's, and 100 µg/ml salmon sperm DNA (ssDNA), followed by washing with 0.3 x SSC/0.1% SDS at 48°C. As a result, several positive clones were obtained. These cDNA clones were separated by performing this plaque hybridization method in a plurality of times. Subsequently, a cDNA fragment of each phage was subcloned into a plasmid vector pBluescript II, and the nucleotide sequence was analyzed using a dideoxy sequence method [Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467]. Based on a putative amino acid sequence, LDL receptors themselves were excluded, and cDNA clones having a sequence very similar to that of LDL receptors were identified. Using these clones as cDNA probes, the cDNA library was screened to thereby obtain overlapping two clones. These were employed as new probes and similar procedure was performed, so as to obtain 5 cDNA clones. The DNA nucleotide sequence determined by these cDNA clones are shown as Sequence ID No. 3. The total length of the sequence was 6961 bp. In the open reading frame of 6639 bp (Sequence ID No. 1) which contained a sequence exhibiting high homology with LDL receptors, there existed on the 5' side an ATG codon which was presumably a translation initiating site and a successive highly hydrophobic sequence consisting of about 30 amino acids. Accordingly, the obtained cDNA was considered to contain the entirety of its length. A putative amino acid sequence is shown as Sequence ID No. 2. The protein consisted of 2213 amino acids. Comparison of the amino acid sequence of the protein with other amino acid sequence data registered at the Genebank, there was a very high similarity to LDL receptors. That is, amino acids 700 - 1,100 in the sequence were very similar to the EGF precursor homology region of LDL receptors, and amino acids 1,100 - 1,640 were also very similar to the ligand binding region of LDL receptors. When the amino acid sequence of the subject protein was compared with other lipoprotein receptor LRP, gp330, and VLDL receptors, similarity was not as high as that observed for LDL receptors. On the C-terminal side of the amino acid sequence of the protein, there was found a highly hydrophobic region which was very similar to the transmembrane region of LDL receptors.

25 Example 3:

30 From liver tissue and brain tissue of a male Japanese white rabbit, intact RNA was extracted through a guanidium thiocyanate/cesium chloride method. The obtained intact RNA was subjected to an oligo (dT) cellulose column method to thereby obtain purified poly(A)⁺RNA. The poly(A)⁺RNA specimens (10 µg each) was modified via a glyoxal method, electrophoresed on 1% agarose gel, and transferred onto a nylon membrane.

35 For human tissue mRNA, commercially available nylon membranes blotted with human tissue poly(A)⁺RNA from various sources were used.

40 Using as a probe part of a ³²P-labelled rabbit cDNA of the present invention, hybridization was performed at 42°C using 50% (rabbit) or 40% (human) formamide, 0.1% SDS, 50 mM phosphate buffer, 5 x Denhardt's, 5 x SSC, and 200 µg/ml of ssDNA, followed by washing with 0.1 x SSC and 0.1% SDS at 50°C. Autoradiography was performed at -70°C for 2 days in the presence of intensifying screen. As a result, in both rabbit liver tissue and brain tissue, mRNA of about 7 kb was detected as well as mRNA of about 15 kb which was considered to result from alternative splicing or polyadenylation. The size of the mRNA of about 7 kb coincided with that of the rabbit cDNA of the present invention. Also, in human liver tissue and brain tissue, it was confirmed that mRNA having the same size was expressed.

Example 4:

Screening of human brain cDNA library for positive clones and determination of the nucleotide sequence of cDNA fragments

5 The human brain cDNA library used in this Example was a commercially obtained cDNA library which was constructed using λ gt10 as a vector. Using partial cDNA of the present invention as a probe, screening of the cDNA library (300,000 plaques) was performed using a plaque hybridization method. Procedures of screening, cloning, and sequencing were as described in Example 2 of the present invention.

10 As a result of screening of the human brain cDNA library, positive clones containing a DNA fragment of about 3 kb were obtained. Analysis of the nucleotide sequence of part of the cDNA fragment revealed that the fragment was highly homologous to the cDNA of the present invention (Sequence ID No. 4).

Example 5:

15 Cloning of cDNA of receptors in the human LDL receptor family:

A human brain cDNA library was subjected to screening using fragments of the cDNA of the present invention and fragments of the cDNA obtained in Example 4 as probes. Procedures of screening, cloning, and sequencing were as 20 described in Example 2 of the present invention.

25 Through screening of the human brain cDNA library, two positive clones containing cDNA fragments of about 6 kb and about 3 kb were obtained. When their nucleotide sequence was analyzed, they were identified to be a cDNA clone containing the cDNA nucleotide sequence obtained in Example 4 and a cDNA clone that overlapped therewith. Using part of these cDNAs as probes, procedures similar to those as described above were performed, to thereby obtain another cDNA clone. The DNA nucleotide sequence indicated by these cDNA clones are shown as Sequence ID No. 7. The total length of the sequence was 6,843 bp. There was an open reading frame having a size of 6,642 bp (Sequence ID No. 5). A putative amino acid sequence is shown as Sequence ID No. 6. The protein consisted of 2,214 amino acids. Comparison of the amino acid sequence with that of rabbit protein shown by Sequence ID No. 2 revealed high homology of not less than 94%.

Example 6

30 Creation of cells that express receptors in the rabbit LDL receptor family:

35 The cDNA as shown by Sequence ID No. 3 was ligated to phosphorylated EcoRI linker pd (CCGAATTCGG) by the use of a T4 DNA ligase, and the resultant ligated product was digested with EcoRI. Separately, a vector for expression, pBK-CMV was digested with EcoRI. The aforementioned DNA was ligated to the EcoRI-digested site of the vector using a T4 DNA ligase.

40 Using the resultant recombinant expression vector in a calcium phosphate method [Chen, C. and H. Okayama (1987) Mol Cell Biol 7: 2745-2752], host cells (CHO-1d1A7) were transformed. The resultant transformants were incubated in a Ham's F-12 selective medium supplemented with 500 μ g/ml of G418, and viable cells were separated as LDL receptor analog protein-expressing cells. The cells were incubated further in the aforementioned medium.

Example 7

45 Ligand analysis of the LDL receptor analog protein by ligand blotting:

50 The obtained LDL receptor analog protein-expressing cells and control cells were suspended in a buffer solution containing 200 mM Tris-maleic acid (pH 6.5), 2 mM calcium chloride, 0.5 mM PMSF, 2.5 μ M leupeptin, and 1% Triton X-100, to thereby solubilize the membrane protein. Solubilized membrane protein fractions were obtained through centrifugation, and electrophoresed by a 4.5-18% gradient SDS-PAGE. Thereafter, the protein was transferred onto a nitrocellulose membrane.

55 Incubation was performed in a buffer of 50 mM Tris-HCl (pH 8.0) containing 125 I-labelled β -VLDL (10 μ g/ml), 2 mM calcium chloride, and 5% bovine serum albumin. Autoradiography was performed at room temperature.

A single band of about 250 kDa was detected in membrane protein fractions prepared using the present protein-expressing cells. This size coincided well with the molecular weight of 248 kDa calculated regarding the amino acid sequence (Sequence ID No. 2) deduced from the cDNA of the present invention. Although a similar band was detected for control cells, the expression level was much lower as compared with the case of the present protein-expressing cells.

Since the protein coded by the cDNA of the present invention is considered to be a novel LDL receptor family recep-

tor, it is expected that through analyses of this protein, details of lipoprotein metabolism mediated by the membrane receptor will be elucidated, and pathology of abnormal lipid metabolism which triggers onset and progress of arteriosclerosis will be clarified.

5 **Sequence ID No. 1**

Length of the Sequence: 6639

Type: nucleic acid

10 Strandedness: double

Topology: linear

15 Molecular type: cDNA to mRNA

Sequence:

ATGGCGACAC	GGAGCAGCAG	GAGGGACTCG	CGACTCCCCT	TCCTATTAC	CCTGGTCCCG	60
20 CTGCTGCCGC	CCGGGGCTCT	CTGGGAGGTG	TGGACGGGGA	CACTGCACGG	CGGGCGCGCG	120
CCCTTACCCC	AGGAGCGGGG	CTTCCCCGTG	GTGGAGGGCG	ACCCGGCGGA	GCTGGGCGTC	180
25 TGGGAGCGCG	GGGATGCCAG	GGGGGGCGAGC	CGGGGGCGACG	AGAAAGGGCT	CCGGAGGAGA	240
CGGAGCGCTG	CCCTGCAGCC	CGAGCCCCATC	AAGCTGTACG	GACAGGTCA	CCTCAATGAT	300
TCCCACAATC	AGATGGTGGT	GCACCTGGCC	GGAGAGAAAAA	CCAACGTGAT	CGTGGCCCTG	360
30 CCCCGGGACA	GCCTGGCGTT	GGCCAGGGCC	AGGAGCAGTG	ATGTGTACGT	GTCTTATGAC	420
TATGGAAAAT	CATTCAATAA	GATTTCACAG	AAATTGAAC	TCGGCGCGGG	AAATAACACA	480
40 GAGGCTGTCC	TGGCCCAAGTT	CTACCACAGC	CCTGGGACA	ACAAACGCTA	CATCTTCGCA	540
GATGCCTACG	CCCAGTATCT	CTGGATCACG	TTTGACTTCT	GCAACACCAT	CCATGGCTTT	600
TCCATCCCGT	TCCGGGCAGC	TGATCTCCTA	CTCCACAGTA	AGGCCTCCAA	CCTTCTCCTG	660
45 CGCTTCCACA	CCTCTCACCC	CAACAACCA	CTGTGGAAGT	CGGATGATTT	TGGCCAGACC	720
TGGATCATGA	TTCAAGAACCA	CGTGAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATCACAAA	730
50 CCAAACACCA	TCTACATCGA	ACGGCACGAA	CCTTCTGGCT	ACTCCACGGT	TTTCCGAAGT	840
ACAGACTTCT	TCCAGTCCCG	GGAAAACCA	GAAGTGATCT	TGGAGGAAGT	GAGAGACTTT	900
55 CAGCTTCCGG	ACAAGTACAT	GTTTGCTACA	AAGGTGGTGC	ATCTCTTGGG	CAGTCCACTG	960
CAGTCTCTG	TCCAGCTCTG	GGTCTCCTTT	GGCCCGAAGC	CCATGCGGGC	CGCCCAAGTT	1020
GTTACAAGAC	ATCCTATCAA	CGAATATTAC	ATCGCGGATG	CCTCGGAGGA	CCAGCTGTTT	1080
60 GTGTGTGTCA	GTCACAGCAA	CAACCGCACC	AACCTCTACA	TCTCGGAGGC	AGAGGGCTTG	1140
AACTTCTCTC	TGTCCCTGGA	GAACGTGCTC	TACTACACCC	CGGGAGGGCC	CGGCAGTGAC	1200
65 ACCTTGGTGA	GGTACTTTGC	AAATGAACCG	TTTGCTGACT	TCCATCGTGT	GGAAAGGGTTG	1260

5 CAGGGACTCT ACATTGCTAC TCTGATTAAT CGTTCTATGA ATGACGGAGAA CATCAGATCT 1320
 GTCATCACCT TTGACAAACG GGGCACCTGG GAATTCTGC AGGCTCCAGC CTTCACGGGG 1380
 TATGGAGAGA AAATCAACTG TGAGCTGTCC GAGGGCTGTT CCCTCCACCT GGCCCAGCGC 1440
 10 CTCAGCCAGC TGCTCAACCT CCAGCTCCGG AGGATCCCCA TCCCTGCCAA GGAGTCGGCG 1500
 CCTQQCCTCA TCATTGCCAC GGCTCACTG CGAAAGAACT TGGCTACCAA GACAAACGTC 1560
 TACATCTCTA GCAGTGCCTGG AGCCAGGTGG CGAGAGGCAC TTCCCTGGACC TCACTACTAT 1620
 ACATGGGGAG ACCATGGGG CATCATCATG CCCATTGCCA AAGGCATGGA AACCAACGAA 1680
 15 CTGAAGTACA GTACCAACGA AGGGGAGACC TGGAAAGCCT TCACCTTCTC TGAGAAGCCC 1740
 GTGTTTGTGT ATGGGCTCCT CACCGAACCC CCCGACAAAGA CCACCGCTTT CACCATCTTT 1800
 20 GGCTCCAACA AGGAGAACCT GCACAGCTGG CTCATCCTCC AGGTCAATGC CACAGACCCC 1860
 CTGGGGCTTC CTTGCACAGA GAACGACTAC AAGCTCTGTT CACCATCTGA TGAGGGGGGG 1920
 AATGAGTGTGTT TGCTTGGACA CAAGACTGTT TTCAAACGGA GGACCCCGCA CGCCACATGC 1980
 25 TTTAACGGAG AAGACTTTGA CAGGCCGGTG GTTGTCTCCA ACTGCTCCTG CACCCGGGAC 2040
 GACTATGAGT GTGACTTTGG CTTCCGGATG AGTGAAGACT TGGCATTAGA GGTGTGTGTT 2100
 CCAGATCCAG GATTTCTGG AAAGTCCTCC CCTCCAGTGC CTTGTCCCCTG GGCTCTACCG 2160
 30 TACAGGCCAT CAAGACGCTA CCCGAAGATT TCTGGGGACA CCTGTAGTGG AGGAGATGTT 2220
 GAGGCACGGC TAGAAGGAGA GCTGGTCCCC TGTCCCCCTGG CAGAAGAGAA CGAGTTCATC 2280
 CTGTACGCCA CCCGCAAGTC CATCCACCGC TATGACCTGG CTTCCGGAAC CACGGACCG 2340
 35 TTGCCCCCTCA CTGGGTTGGC GCCAGCAGTG CCCCTGGACT TTGACTATGA GCACAACTGC 2400
 CTCTATTGGT CTGACCTGGC CTTGGACGTC ATCCAGGCC TCTGTTGAA CGGGACTACA 2460
 GGACAAGAGG TGATCATCAA CTCTGACCTG GAGACGGTAG AAGCTTGCG TTTTGAACCC 2520
 40 CTCAGCCAAT TACTTTACTG GGTGGACGCA GGCTTTAAAAA ACATGGAGGT AGCCAATCCA 2580
 GATGGTCACT TCCGACTCAC CGTCGTCAAT TCCCTGGTGC TGGATCGGCC CGGGGCCCTG 2640
 45 GTCCCTTGTGC CCCAAGAAGG GATCATGTTC TGGACCGACT GGGGAGACCT GAAGCCTGGG 2700
 ATTTATCGGA GCAACATGGA CGGATCTGCC CCCTATCGCC TCGTGTGGAA GGATGTGAAG 2760
 TCCCCCAATG GCATTTCCGT GGACGGATCAG TGGATCTACT GGACGGATGCC CTACCTGGAC 2820
 50 TGCATTGAGC CCATCACGTT CAGCGGCCAG CAGCGCTCCG TCATCCTGGA CAGACTCCCG 2880
 CACCCCTATG CCATTGCTGT CTTTAAGAAT GAGATTACT GGGATGACTG GTCACAGCTC 2940

5	ACCATATTCC CAGCTTCTAA GTACAGCGGG TCCCAGATCC ACATTCTCCC CAGCCAGCTC	3000
	ACGGGGCTGA TGGACATGAA GATCTTCTAC AAGGGGAAGA ACACAGGAAG CAATGCCGTG	3060
	GTACCCAGGC CGTSCAGCCT GCTGTGCCCTG CCCAGAGCCA ACAACACCAA AAGCTGCAGG	3120
10	TGTCCAGATG GCGTGGCCAG CAGTGTCTC CCTTCCGGGG ACCTGATGTG TCACTGCCCT	3180
	AAGGGCTACG ACCTGAAGAA CAACACGTCT GTCAAAGAAG AAGACACCTG TCTGCCCAAC	3240
	CAGTACCGCT CCAGCAACGG GAACTGCATC AACAGCATCT GGTGGTGCGA TTTGCACAAAC	3300
	GAATGCGGAG ACATGAGCGA CGAGAAGAAC TGCCCTACCA CCATCTGCAG CCTGGACACC	3360
15	CAGTCCGTT GCCAGGAGTC TGGGACGTGC ATCCCGCTCT CCTACAAATG TGACCTCCAG	3420
	GATGACTGTG GGGACAACAG TGACGAAAGG CACTGTGAAA TGCACCAGTG CCGGAGCCAC	3480
	GAATACAACG GCAGCTCGGG CATGTGCATC CGCTCTCCT GGGTGTGCAG CGGGGACAAC	3540
20	GAATGCGGAG ACCTGGTCCGA CGAGGCAAC TCCACACCCA TCTATCACAC CTGTGAGGCC	3600
	TCCAACATTCC AGTGCCGCAA CGGGCACTGC ATCCCCCAGC GGTGGCCGTG TGACGGCCAC	3660
25	GCCGACTGCC AGGATGGCTC TGATGAGGAT CCAGCCAAT GTGAGAAGAA GTGCAACGGC	3720
	TTCCGCTGCC CGAACGGCAC CTGCATTCCC TCCACCAAGC ACTGTGACGG CCTGCACGAT	3780
	TGCTCCGACG GCTCCGACGA GCAGCACTGC GAGCCCCCTGT GTACACGGTT CATGGACTTC	3840
30	GTGTGTAAGA ACCGCCAGCA GTGCCCTTC CACTCCATGG TGTGGATGG GATCATCCAG	3900
	TGCCGTGACG GCTCCGACGA GGACCCAGCC TTTGCAGGAT GCTCCCGAGA CCCCCGAGTTC	3960
	CACAAGGTGT CGCATGAGTT CGGCTTCAG TCTCAGAACG GCCTGTGCAT CAGCTTGATC	4020
35	TGGAAGTGCG ACGGGATGCA TGACTCCGG GACTACTCCG ACCAGGCCAA CTGTAAAAC	4080
	CCCACAGAACG CCCCCAACTG CTCCCCCTAC TTCCAGTTCC GCTGTGACAA TGGCCACTCC	4140
	ATCCCCAACG GGTGGAAGTG TGACAGGGAG AATGACTGTG GGGACTGGTC CGACCGAGAAC	4200
40	GAATGCGGAG ATTACATGT ACTTCCCTCT ACCACTCCG CACCCCTCCAC GTGTCTGCC	4260
	AATTACTACC GCTGCCGGCCG GGGGGCTTCG GTGATAGACA CGTGGGTTTG TGACCCGTAC	4320
45	CGAGATGGCG CAGATGGATC CGACGGAGAA GCCTCCCCCT CGCTCCCCAA TGTCACTGCC	4380
	ACCTCCCTCC CTCCTCCAGCC TGGACCATGC GACCGATTG AGTTTGACTG CCACCCAGCA	4440
	AAGAAAGTCCA TCCCTAACTG GAGACGCTCT GACGGCCATC AGGATTCCCA GGATGCCAG	4500
50	GACGAGGCCA ACTGCCAC TCACAGCACC TTGACCTGCA TGAGCTGGGA GTTCAAGTGT	4560
	GAGGATGGCG AGGGCTGCAT CGTCTCTCA CAACGCTGGG ACGGCTTCCT GGACTGCTCA	4620

55

5 GATGAGAGCC ACGAGAAGGC CTGCAGTGAT GAGTTAACTG TATACAAAGT ACAGAATCTT 4680
 CAGTGGACAG CTGACTTCTC TGGGAATGTC ACTTTGACCT GGATGGGGCC CAAAAAAATG 4740
 CCCTCTGCTG CTTGTGTATA CAACGTGTAC TATAGAGTTG TTGGAGAGAG CATATGGAAG 4800
 ACTCTGGAGA CTCACACCAA TAAGACAAAC ACTGTATTAA AAGTGTGAA ACCAGATACC 4860
 10 ACCTACCAGG TTAAAGTCCA GGTTCACTGTC CTGAGCAAGG TGCAACACAC CAATGACTTT 4920
 GTCACCTTGA CAACTCCAGA GGGATTGCCA GACGCCCCCTC AGAACCTCCA GCTGTCGCTC 4980
 CACGGGGAAG AGCAAGGTGT GATTGTGGGC CACTGGAGCC CTCCCCACCCA CACCCACGCC 5040
 15 CTCATTCGCG AATACATTGT AGAGTATAGC AGGAGTGGTT CCAAGGTGTG GACTTCAGAA 5100
 AGGGCTGCTA GTAACTTTAC AGAAATAAG AACTTGTGG TCAACACCCCT GTACACCGCTC 5160
 20 ACAGTCCCTG CGGTGACCGAG TCGTGGATA GGAAACTGGA GCGATTCCAA ATCCATTAC 5220
 ACCGTGAAAG GAAAAGCGAT CCCGCCACCA AATATCCACA TTGACAACTA CGATQAAAAT 5280
 TCCCTGAGTT TTACCCCTCAC CGTGGATGGG AACATCAAGG TGAATGGCTA TGTGGTGAAC 5340
 25 CTTTTCTGGG CATTGACAC CCACAAACAA GAGAAGAAAA CCATGAACTT CCAAGGGACC 5400
 TCAGTGTCCC ACAAAAGTTGG CAATCTGACA GCACAGACGG CCTATGAGAT TTCCCGCTGG 5460
 CCCAAGACTG ACTTGGGCGA TAGTCCTCTG TCATTTGAGC ATGTCACCGAC CAGAGGGTT 5520
 30 CGCCCACCTG CTCCCTAGCCT CAAGGCCAGG GCTATCAATC AGACTGCAGT GGAATGCCACC 5580
 TGGACAGGCC CCAGGAATGT GGTGTATGCC ATTTTCTATG CCACATCCTT CCTGGACCTC 5640
 TACCGCAACC CAAGCAGCCT GACCACGCCG CTGCACAAAC CAACCGTCT CCGCGGTAAG 5700
 35 CATGAGCAGT ATCTGTTCT GGTCCGGGTG GTGATGCCCT ACCAAGGGCC GTCCTCCGAC 5760
 TACGTGGTCG TGAAGATGAT CCCCCGACAGC AGGCTTCCTC CCCCCGACCT GCATCCCCGT 5820
 40 CACACGGCA AGACCTCGGC CGTCATCAAG TGGGAGTCGC CCTACGACTC TCCTGACCAG 5880
 GACCTGTTCT ATGCGATCGC AGTTAAAGAT CTGATACGAA AGACGGACCG GAGCTACAAA 5940
 CTCAAGTCCC GCAACAGCAC CGTGGAGTAC ACCCTGAGCA AGCTGGAGCC CGGACGGAAA 6000
 45 TACCAAGTCA TTGTGCAGCT GGGGAACATG ACCAAAGATG CCAGTGTGAA GATCACCAC 6060
 GTTTCGTTAT CGGCACCCGA TGCCTAAAAA ATCATAACAG AAAATGACCA CGTCCTTCTC 6120
 TTCTGGAAAAA GTCTAGCTCT AAAGGAAAAG TATTTTAACG AAAGCAGGGG CTACCGAGATA 6180
 50 CACATTTTG ATAGGCCAT GAATATCACC CCATACCTTG CGAATACTAC TGACAATTTC 6240
 TTTAAAATTT CCAACCTGAA GATGGGTAC AATTACACAT TCACGGTCCA CGCACCGATGC 6300

5 CTTTTGGGCA GCCAGATCTG CGGGGAGCCT GCCGTGCTAC TCTATGATGA CCTGGGGTCT 6360
 GGTGGCGATG CGTCGGCGAT GCAGGCTGCC AGGTCTACTG ATGTCGGGCC CGTGGTGGTG 6420
 CCCATCCTGT TTCTGATACT GCTGAGCCTG GGGGTCGGGT TTGCCATCCT GTACACCAAG 6480
 10 CATCCGAGGC TGCAGAGCAG CTTCACCCGC TTGCGCAACA CCCACTACAG CTCCAGACTC 6540
 GGCTCCGCCA TCTTCTCCTC TCGGGATGAC TTGGGGGAGG ATGATGAACA TGCTCCTATG 6600
 ATCACTGGAT TTTCGGACGA CGTCCCCATG GTGATAGCC 6639

Sequence ID No. 2

15 Length of the Sequence: 2213

Type: amino acid

Topology: linear

20 Molecular type: Protein

Sequence:

25 Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe

5 10 15

Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr

30 20 25 30

Arg Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe

35 40 45

45 Arg Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly

50 55 60

Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg

40 65 70 75 80

Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val

85 90 95

45 Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu

100 105 110

50 Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala

115 120 125

55

Arg Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
 130 135 140
 5 Phe Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr
 145 150 155 160
 10 Glu Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
 165 170 175
 Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
 180 185 190
 15 Phe Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
 195 200 205
 20 Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
 210 215 220
 25 Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
 225 230 235 240
 Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
 245 250 255
 30 Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
 260 265 270
 35 Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
 275 280 285
 40 Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp
 290 295 300
 Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu
 305 310 315 320
 45 Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
 325 330 335
 50 Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
 340 345 350

55

Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
 355 360 365
 5 Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
 370 375 380
 10 Ser Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp
 385 390 395 400
 15 Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
 405 410 415
 Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
 420 425 430
 20 Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
 435 440 445
 25 Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
 450 455 460
 Ile Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg
 465 470 475 480
 30 Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
 485 490 495
 35 Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
 500 505 510
 40 Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
 515 520 525
 45 Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
 530 535 540
 His Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
 545 550 555 560
 50 Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe
 565 570 575
 55

Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
 5 580 585 590
 Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
 10 595 600 605
 Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
 15 610 615 620
 Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
 625 630 635 640
 Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
 20 645 650 655
 His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
 660 665 670
 Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
 25 675 680 685
 Arg Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly
 30 690 695 700
 Phe Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
 35 705 710 715 720
 Tyr Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
 725 730 735
 Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro
 40 740 745 750
 Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile
 45 755 760 765
 His Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr
 770 775 780
 Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
 50 785 790 795 800
 55

Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
 805 810 815
 Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr
 820 825 830
 Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
 835 840 845
 Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
 850 855 860
 Arg Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
 865 870 875 880
 Val Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp
 885 890 895
 Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
 900 905 910
 Arg Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp
 915 920 925
 Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg
 930 935 940
 Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro
 945 950 955 960
 His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp
 965 970 975
 Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln
 980 985 990
 Met Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile
 995 1000 1005
 Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro
 1010 1015 1020

5 Cys Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg
 1025 1030 1035 1040
 Cys Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met
 1045 1050 1055
 10 Cys Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys
 1060 1065 1070
 Glu Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn
 1075 1080 1085
 Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp
 1090 1095 1100
 20 Met Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr
 1105 1110 1115 1120
 25 Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
 1125 1130 1135
 Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys
 30 1140 1145 1150
 Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
 1155 1160 1165
 35 Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
 1170 1175 1180
 40 Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
 1185 1190 1195 1200
 Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
 45 1205 1210 1215
 Cys Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala
 1220 1225 1230
 50 Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
 1235 1240 1245

55

Ile Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly
 1250 1255 1260
 5 Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe
 1265 1270 1275 1280
 10 Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
 1285 1290 1295
 Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala
 1300 1305 1310
 15 Gly Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
 1315 1320 1325
 20 Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
 1330 1335 1340
 Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
 25 1345 1350 1355 1360
 Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp
 1365 1370 1375
 30 Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
 1380 1385 1390
 Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu
 35 1395 1400 1405
 Pro Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
 1410 1415 1420
 40 Cys Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr
 1425 1430 1435 1440
 45 Arg Asp Cys Ala Asp Cys Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro
 1445 1450 1455
 Asn Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg
 50 1460 1465 1470

55

Phe Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg
 1475 1480 1485
 5 Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn
 1490 1495 1500
 10 Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys
 1505 1510 1515 1520
 15 Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
 1525 1530 1535
 20 Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
 1540 1545 1550
 25 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
 1555 1560 1565
 30 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala
 1570 1575 1580
 35 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
 1585 1590 1595 1600
 40 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
 1605 1610 1615
 45 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
 1620 1625 1630
 50 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
 1635 1640 1645
 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu
 1650 1655 1660
 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly
 1665 1670 1675 1680
 55 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val
 1685 1690 1695

55

Trp Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
 1700 1705 1710
 5 Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg
 1715 1720 1725
 10 Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly
 1730 1735 1740
 Lys Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn
 1745 1750 1755 1760
 15 Ser Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly
 1765 1770 1775
 20 Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys
 1780 1785 1790
 25 Lys Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn
 1795 1800 1805
 Leu Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp
 30 1810 1815 1820
 Leu Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val
 1825 1830 1835 1840
 35 Arg Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala
 1845 1850 1855
 Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe
 40 1860 1865 1870
 Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr
 45 1875 1880 1885
 Thr Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr
 1890 1895 1900
 50 Leu Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp
 1905 1910 1915 1920
 55

Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His
 1925 1930 1935
 5 Leu His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu
 1940 1945 1950
 10 Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val
 1955 1960 1965
 Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg
 1970 1975 1980
 15 Asn Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys
 1985 1990 1995 2000
 20 Tyr His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val
 2005 2010 2015
 25 Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile
 2020 2025 2030
 Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys
 30 2035 2040 2045
 Glu Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp
 2050 2055 2060
 35 Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe
 2065 2070 2075 2080
 40 Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val
 2085 2090 2095
 Gln Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val
 45 2100 2105 2110
 Leu Leu Tyr Asp Glu Leu Gly Ser Gly Asp Ala Ser Ala Met Gln
 2115 2120 2125
 50 Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Pro Ile Leu Phe
 2130 2135 2140

55

5 Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys
 2145 2150 2155 2160
 His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr
 2165 2170 2175
 10 Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly
 2180 2185 2190
 Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val
 15 2195 2200 2205
 Pro Met Val Ile Ala
 2210
 20 Sequence ID No. 3
 Length of the Sequence: 6961
 25 Type: nucleic acid
 Strandedness: double
 Topology: linear
 30 Molecular type: cDNA to mRNA
 Feature:
 Name/Key: sig peptide
 35 Location: 178..261
 Identification method: S
 Name/Key: mat peptide
 40 Location: 262..6816
 Identification method: S
 Sequence:
 45 CCGCGAGCCG CACACGTGAC GGGGCGCGCG CGCGCGCGCG CGCGGCCGAGC GGGACCCAGC 60
 GGCTGCCCGG AGCCCCGGGA GCGGGCGCGCG CGCGGCCCGG GCGCCGCGCG TCGGCCGGCG 120
 50 GCGCCGCTGCA CATTCTCTCC TGGCGGGCGGC GCCACCTGCA GCCGCGTTCG CCCGAACATG 180
 Met

55

1

5 GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCC TTC CTA TTC ACC 228
 Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr
 10 5 10 15
 15 CTG GTC GCG CTG CTG CCG CCC GGG GCT CTC TGC GAG GTG TGG ACG CGG 276
 Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Arg
 20 20 25 30
 25 ACA CTG CAC GGC GCC CGC GCG CCC TTA CCC CAG GAG CGG GCC TTC CGC 324
 Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe Arg
 30 35 40 45
 35 GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GAG CGC GGG GAT 372
 Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly Asp
 40 50 55 60 65
 45 50 CCC AGG CGG CGG ACC CGG CGG GAC GAG AAG CGG CTC CGG AGG AGA CGG 420
 Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg Arg
 50 55 60 65
 55 AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG CTG TAC GGA CAG CTC AGC 468
 Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser
 60 65 70 75 80
 65 70 75 80
 70 CTC AAT GAT TCC CAC AAT CAG ATC GTG GTG CAC TGG GCC GGA GAG AAA 516
 Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys
 75 80 85 90 95
 80 85 90 95
 85 ACC AAC GTG ATC GTG GCC TTG CCC CGG GAC AGC CTG GCG TTG GCC AGG 564
 Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg
 90 95 100 105 110
 95 100 105 110
 100 ACC AGG AGC AGT GAT GTG TAC GTG TCT TAT GAC TAT GGA AAA TCA TTC 612
 Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe
 105 110 115 120 125
 110 115 120 125
 115 120 125 130 135
 120 125 130 135 140
 125 130 135 140 145
 130 135 140 145

55

	AAT AAG ATT TCA GAG AAA TTG AAC TTC GGC GCG GGA AAT AAC ACA GAG	660		
5	Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Tyr Glu			
	150	155	160	
	GCT GTG GTG GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAA CGG TAC	708		
10	Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr			
	165	170	175	
	ATC TTC GCA GAT GCC TAC CCC CAG TAT CTC TGG ATC ACG TTT GAC TTC	756		
15	Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe			
	180	185	190	
	TGC AAC ACC ATC CAT GGC TTT TCC ATC CCC TTC CGG GCA CCT GAT CTC	304		
20	Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu			
	195	200	205	
	CTA CTC CAC AGT AAG GCC TCC AAC CTT CTC CTG GGC TTC GAC AGG TCT	852		
	Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser			
25	210	215	220	225
	CAC CCC AAC AAG CAG CTG TGG AAG TCG GAT GAT TTT GGC CAG ACC TGG	900		
	His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp			
	230	235	240	
30	ATC ATG ATT CAA GAA CAC GTG AAG TCC TTT TCT TGG GGA ATT GAT CCC	948		
	Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro			
	245	250	255	
35	TAT GAC AAA CCA AAC ACC ATC TAC ATC GAA CGG CAC GAA CCT TCT GGC	996		
	Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly			
40	260	265	270	
	TAC TCC ACG GTT TTC CGA ACT ACA GAC TTC TTC CAG TCC CGG GAA AAC	1044		
	Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn			
45	275	280	285	
	CAG GAA GTG ATC TTG GAG GAA GTG AGA GAC TTT CAG CTT CGG GAC AAG	1092		

55

5 Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys
 290 295 300 305
 TAC ATG TTT GCT ACA AAG CTG CTG CAT CTC TTG CCC ACT CCA CTG CAG 1140
 Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu Gln
 10 310 315 320
 TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG CGG GGC 1188
 Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala
 15 325 330 335
 GCC CAG TTT GTT ACA AGA CAT CCT ATC AAC GAA TAT TAC ATC GCG GAT 1236
 Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp
 20 340 345 350
 GCC TCG CAG GAC CAG CTG TTT GTG TGT GTC ACT CAC AGC AAC AAC CGC 1284
 Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg
 25 355 360 365
 ACC AAC CTC TAC ATC TCG GAG GCA GAG GGC TTG AAG TTC TCT CTG TCC 1332
 Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser
 30 370 375 380 385
 CTG GAG AAC CTG CTC TAC TAC ACC CCG CGA GGG GCC GGC AGT GAC ACC 1380
 35 Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp Thr
 390 395 400
 TTC CTG ACC TAC TTT GCA AAT GAA CCG TTT GCT GAC TTC CAT CGT CTG 1428
 40 Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val
 405 410 415
 GAA GGG TTG CAG GGA GTC TAC ATT CCT ACT CTG ATT AAT CGT TCT ATG 1476
 Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met
 420 425 430
 45 AAT GAG GAG AAC ATG AGA TCT GTC ATC ACC TTT GAC AAA CGG CCC ACC 1524
 Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Thr

55

	435	440	445	
5	TGG GAA TTT CTG CAG GCT CCA GCC TTC ACG CGG TAT CCA GAC AAA ATC			1572
	Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile			
10	450	455	460	465
	AAC TGT GAG CTG TCC GAG GGC TGT TCC CTC CAC CTG GCC CAG CGC CTC			1620
	Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg Leu			
	470	475	480	
15	AGC CAG CTG CTC AAC CTC CAG CTC CGG AGG ATG CCC ATC CTG TCC AAG			1668
	Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys			
20	485	490	495	
	GAG TCG GCG CCT GGC CTC ATC ATT GCC ACG CGC TCA GTG GGA AAG AAC			1716
	Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn			
25	500	505	510	
	TTC GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG			1764
	Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ala Gly Ala Arg			
30	515	520	525	
	TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAT ACA TGG CGA GAC CAT			1812
	Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His			
35	530	535	540	545
	GGC GGC ATC ATC ATG GCC ATT GCC CAA GGC ATG GAA ACC AAC GAA CTG			1860
	Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu			
40	550	555	560	
	AAG TAC AGT ACC AAC GAA GGG GAG ACC TGG AAA GCC TTC ACC TTC TCT			1908
45	Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe Ser			
	565	570	575	
	GAG AAG CCC GTG TTT GTG TAT GGG CTC CTC ACG GAA CCC GGC GAG AAG			1956
50	Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys			
	580	585	590	

EP 0 773 290 A2

AGC ACC GTC TTC ACC ATC TTT GGC TCC AAC AAG GAG AAC GTC CAC AGC 2004
Ser. Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser
5 595 600 605
TGG CTC ATC CTC CAG GTC AAT GCC ACA GAC GCC CTG GGG GTT CCT TGC 2052
10 Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys
610 615 620 625
ACA GAG AAC GAC TAC AAG CTC TGG TCA CCA TCT GAT GAG CGG CGG AAT 2100
15 Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn
630 635 640
GAG TGT TTG CTT GGA CAC AAG ACT GTT TTC AAA CGG AGG ACC CCG CAC 2148
20 Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro H.s
645 650 655
GCC ACA TGC TTT AAC GGA GAA GAC TTT GAC AGG CCG GTG GTT GTG TCC 2196
25 Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Ser
660 665 670
AAC TGC TCC TCC ACC CCG GAG GAC TAT GAG TGT GAC TTT GGC TTC CCG 2244
30 Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Arg
675 680 685
ATG ACT CAA GAC TTG GCA TTA GAG GTC TGT GTT CCA GAT CCA CGA TTT 2292
35 Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly Phe
690 695 700 705
TCT GGA AAG TCC TCC CCT CCA GTG CCT TGT CCC GTG GGC TCT ACG TAC 2340
40 Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr Tyr
710 715 720
AGG CGA TCA AGA GGC TAC CGG AAG ATT TCT CGG GAC ACC TGT AGT GGA 2388
45 Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser Gly
725 730 735
GGA GAT GTT GAG GCA CGG CTA GAA CGA GAG CTG GTC CCC TGT CCC CTG 2436
50

55

	Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro Leu			
5	740	745	750	
	GCA GAA GAG AAC GAG TTC ATC CTG TAC CCC ACG CGC AAG TCC ATC CAC			2484
	Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile His			
10	755	760	765	
	CGC TAT GAC CTG CCT TCC CGA ACC ACG GAG CAG TTG CCC CTC ACT CGG			2532
	Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr Gly			
15	770	775	780	785
	TTC CGG CGA GCA GTG CCC CTG GAC TTT GAC TAT GAG CAC AAC TGC CTG			2580
	Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys Leu			
20	790	795	800	
	TAT TGG TCT GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG AAC			2628
25	Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu Asn			
	805	810	815	
	CGG ACT ACA CGA CAA GAG GTG ATC ATC AAC TCT GAC CTG GAG ACG GTC			2676
30	Cly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr Val			
	820	825	830	
	CAA CCT TTG GCT TTT GAA CCC CTC AGC CAA TTA CTT TAC TGG GTG GAC			2724
35	Cly Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val Asp			
	835	840	845	
	CCA GGC TTT AAA AAG ATC GAG GTC GCA AAT CCA CAT CGT GAC TTC CGA			2772
40	Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe Arg			
	850	855	860	865
	CTC ACC GTC GTC AAT TCC TCG GTG CTG GAT CGG CCC CGG GCC CTG GTC			2820
45	Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu Val			
	870	875	880	
50	CTT GTG CCC CAA GAA GGG ATC ATG TTC TCG ACC GAC TGG GGA GAC CTG			2868
	Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp Leu			
55				

EP 0 773 290 A2

	885	890	895	
5	AAG CCT GGG ATT TAT CCG AGC AAC ATG GAC GGA TCT GCC GCC TAT CGC			2916
	Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr Arg			
	900	905	910	
10	CTC CTG TCG GAG GAT GTG AAG TGG CCC AAT GCC ATT TCC GTG GAC GAT			2964
	Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp Asp			
	915	920	925	
15	CAG TGG ATC TAC TGG ACC GAT CCC TAC CTG GAC TGC ATT GAG CGC ATC			3012
	Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg Ile			
	930	935	940	945
20	ACG TTC AGC GGC CAG CAG CGC TCC GTC ATC CTG GAC AGA CTC CCG CAC			3060
	Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro His			
	950	955	960	
25	CCC TAT GCC ATT GCT GTC TTT AAG AAT GAG ATT TAC TGG CAT GAC TGG			3108
	Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp Trp			
	965	970	975	
30	TCA CAG CTC AGC ATA TTC CGA GCT TCT AAG TAC AGC GGG TCC CAG ATG			3156
	Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln Met			
	980	985	990	
35	GAG ATT CTG GCC AGC CAG CTC ACG CGG CTG ATG GAC ATG AAG ATC TTC			3204
	Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile Phe			
	995	1000	1005	
40	TAC AAG GGG AAG AAC ACA CGA AGC AAT GCG TGT GTA CCC AGG CCG TCC			3252
	Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro Cys			
	1010	1015	1020	1025
45	AGC CTG CTG TGC CTG CCC AGA GCC AAC AAC AGC AAA AGC TCC AGG TGT			3300
	Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg Cys			
	1030	1035	1040	

55

5	CCA GAT GGC GTG GCC AGC ACT GTC CTC CCT TCC GGG GAC CTG ATG TGT	3348
	Pro Asp Gly Val Ala Ser Ser Val Leu Prc Ser Gly Asp Leu Met Cys	
	1045 1050 1055	
10	GAC TGC CCT AAG GGC TAC GAG CTG AAG AAC AAC ACG TGT GTC AAA GAA	3396
	Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys Glu	
	1060 1065 1070	
15	GAA GAC ACC TGT CTG CGC AAC CAG TAC CGC TGC AGC AAC GGG AAC TGC	3444
	Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys	
	1075 1080 1085	
20	ATC AAC AGC ATC TGG TGG TGC GAT TTC GAC AAC GAC TGC GGA GAC ATG	3492
	Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met	
	1090 1095 1100 1105	
25	AGC GAC GAG AAG AAC TGC CCT ACC ACC ATC TGC GAC CTG GAC ACC CAG	3540
	Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr Gln	
	1110 1115 1120	
30	TTC CGT TGC CAG CAG TCT GGG ACG TGC ATC CCG CTC TCC TAC AAA TCT	3588
	Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys Cys	
	1125 1130 1135	
35	GAC CTC GAG GAT GAC TGT GGG GAC AAC AGT GAC GAA AGG CAC TGT GAA	3636
	Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys Glu	
	1140 1145 1150	
40	ATG CAC CAG TGC CGG AGC GAC GAA TAC AAC TGC AGC TCG GGC ATG TGC	3684
	Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys	
	1155 1160 1165	
45	ATC CGC TCC TCC TGG GTG TGC GAC GGG GAC AAC GAC TGC AGG GAC TGG	3732
	Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp	
50	1170 1175 1180 1185	
	TCC GAC GAG GCC AAC TGC ACA GCC ATC TAT CAC ACC TGT GAG GCC TCC	3780

Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala Ser
 1190 1195 1200
 5 AAC TTC CAG TGC CGC AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG TGT 3828
 Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala Cys
 1205 1210 1215
 10 GAC GGC GAC GCC GAC TGC CAG CAT GGC TCT GAT GAG CAT CCA GCC AAC 3876
 Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala Asn
 1220 1225 1230
 15 TGT GAG AAG AAG TGC AAC GGC TTC CGC TGC CCG AAC GGC ACC TGC ATT 3924
 Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile
 1235 1240 1245
 20 CCC TCC ACC AAG CAC TGT GAC GGC CTC CAC GAT TGC TCG GAC GGC TCC 3972
 Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly Ser
 1250 1255 1260 1265
 25 GAC GAG CAG CAC TGC GAG CCC CTG TGT ACA CGG TTC ATG GAC TCC GTG 4020
 Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe Val
 1270 1275 1280
 30 TGT AAG AAC CGC CAG CAG TGC CTC TCC CAC TCC ATG GTG TGC SAT GCG 4068
 Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp Gly
 1285 1290 1295
 35 ATC ATC CAG TGC CGT GAC GGC TCC GAC GAG GAC CCA GCC TTT GCA GGA 4116
 Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala Gly
 1300 1305 1310
 40 TGC TCC CGA GAC CCC GAG TTC CAC AAG GTG TCC GAT GAG TTC GGC TTC 4164
 Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly Phe
 1315 1320 1325
 45 CAG TGT CAG AAC GGC GTG TGC ATC AGC TTG ATC TGG AAG TGC GAC GGG 4212
 Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp Gly
 50

1330	1335	1340	1345
ATG GAT GAC TGC GGG GAC TAC TCC GAC GAG GCC AAC TGT GAA AAC CCC			4260
Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn Pro			
1350	1355	1360	
ACA GAA GCC CCC AAC TGC TCC CGC TAC TTC CAG TTC CGG TGT GAC AAT			4308
Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp Asn			
1365	1370	1375	
GGC CAC TGC ATC CCC AAC AGG TGG AAG TGT GAC AGG GAG AAT GAC TGT			4356
Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp Cys			
1380	1385	1390	
GGG GAC TGG TCC GAC GAG AAG GAC TGT CGA GAT TCA CAT GTA CTT CCG			4404
Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu Pro			
1395	1400	1405	
TCT ACG ACT CCT GCA CCC TCC ACC TGT CTG CCC AAT TAC TAC CGC TCC			4452
Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg Cys			
1410	1415	1420	1425
GGC CGG CGG GCC TGC GTG ATA GAC ACG TGG GTT TGT GAC GGG TAC CGA			4500
Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr Arg			
1430	1435	1440	
GAT TGC GCA GAT GGA TCC GAC GAG GAA GCC TGC CCC TCG CTC CCC AAT			4548
Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro Asn			
1445	1450	1455	
GTC ACT GCC ACC TCC TCC CCC TCC CAG CCT GGA CGA TGC GAC CGA TTT			4596
Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg Phe			
1460	1465	1470	
GAC TTT GAG TCC CAC CAG CCA AAG AAG TGC ATC CCT AAC TGG AGA CGC			4644
Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg Arg			
1775	1480	1485	

55

EP 0 773 290 A2

	TGT GAC GGC CAT CAG GAT TGC CAG GAT GGC CAG GAC GAG GCC AAC TGC	4692
5	Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn Cys	
	1490 1495 1500 1505	
	CCC ACT CAC AGC ACC TTG ACC TGC ATG AGC TGG GAG TTC AAG TGT GAG	4740
10	Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys Glu	
	1510 1515 1520	
	GAT GGC GAG GCC TGC ATC GTG CTG TCA GAA CGC TGC GAC GGC TTC CTG	4788
15	Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe Leu	
	1525 1530 1535	
	GAC TGC TCA GAT GAG AGC GAC GAG AAG CCC TGC ACT GAT GAG TTA ACT	4836
20	Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu Thr	
	1540 1545 1550	
	GTA TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG AAT	4884
25	Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly Asn	
	1555 1560 1565	
	GTC ACT TTG ACC TGG ATG CGG CCC AAA AAA ATG CCC TCT GCT GCT TGT	4932
30	Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala Cys	
	1570 1575 1580 1585	
	GTA TAC AAC GTG TAC TAT AGA GTT GTT GGA GAG AGC ATA TGG AAG ACT	4980
35	Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys Thr	
	1590 1595 1600	
40	CTG GAG ACT CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTG TTG AAA	5028
	Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu Lys	
	1605 1610 1615	
	CCA GAT ACC ACC TAC CAG GTT AAA GTG CAG GTT CAG TGC CTG AGC AAG	5076
45	Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser Lys	
	1620 1625 1630	
50	GTG CAC AAC ACC AAT GAC TTT GTG ACC TTG AGA ACT CCA GAG GGA TTG	5124

55

Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly Leu
 1635 1640 1645
 5 CCA GAC GCC CCT CAG AAC CTC CAG CTG TCG CTC CAC CGG GAA GAG GAA 5172
 Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu
 1650 1655 1660 1665
 10 GGT GTG ATT GTG GGC CAC TGG AGC CCT CCC ACC CAC ACC CAC GGC CTC 5220
 Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly Leu
 1670 1675 1680
 15 ATT CGC GAA TAC ATT GTC GAG TAT AGC AGG ACT GGT TCC AAG GTG TGG 5268
 Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val Trp
 1685 1690 1695
 20 ACT TCA GAA AGG GCT GCT ACT AAC TTT ACA GAA ATA AAG AAC TTG TTG 5316
 Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu Leu
 1700 1705 1710
 25 GTC AAC ACC CTG TAC ACC GTC AGA GTG GCT GCG GTG ACG AGT CGT GGG 5364
 Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg Gly
 1715 1720 1725
 30 ATA GGA AAC TGG ACC GAT TCC AAA TCC ATT ACC ACC CTG AAA GGA AAA 5412
 Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly Lys
 1730 1735 1740 1745
 35 GCG ATC CCG CCA CCA AAT ATC CAC ATT CAC AAC TAC GAT GAA AAT TCC 5460
 Ala Ile Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn Ser
 1750 1755 1760
 40 CTG AGT TTT ACC CTG ACC GTG GAT GGG AAC ATC AAG GTC AAT CGC TAT 5508
 Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly Tyr
 1765 1770 1775
 45 GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAA CAA GAG AAG AAA 5556
 Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys Lys
 50
 55

	1780	1785	1790	
5	ACC ATG AAC TTC CAA GGG AGC TCA GTG TCC CAC AAA CTT CCC AAT CTG			5604
	Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn Leu			
10	1795	1800	1805	
	ACA GCA CAG ACG GCG TAT GAG ATT TCC GCC TGG GCC AAG ACT GAC TTG			5652
	Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp Leu			
15	1810	1815	1820	1825
	GCG GAT AGT CCT CTG TCA TTT GAG CAT GTC ACG ACC AGA GGG GTT CGC			5700
	Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val Arg			
20	1830	1835	1840	
	CCA CCT GCT CCT AGC CTC AAG GCC AGG GCT ATC AAT CAG ACT GCA GTG			5748
	Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala Val			
25	1845	1850	1855	
	GAA TGC ACC TGG ACA GGC CCC AGG AAT CTG GTG TAT GGC ATT TTC TAT			5796
	Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe Tyr			
30	1860	1865	1870	
	GCC ACA TCC TTC CTG GAC CTC TAC CGC AAC CCA AGC AGC CTG ACC ACC			5844
	Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr Thr			
35	1875	1880	1885	
	CCG CTG CAC AAC GCA ACC GTG CTC GTC GGT AAG GAT GAG CAG TAT CTG			5892
	Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr Leu			
40	1890	1895	1900	1905
	TTT CTG GTC CGG GTG ATG CCC TAC CAA GGG CCG TCC TCG GAC TAC			5940
	Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp Tyr			
45	1910	1915	1920	
	CTG GTC CTG AAG ATG ATC CCG GAC ACC AGG CTT CCT CCC CGG CAC CTG			5988
	Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His Leu			
50	1925	1930	1935	

	CAT GCC CTT CAC ACC GGC AAG ACC TCG GCC CTC ATC AAG TGG GAG TCG	6036		
	His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu Ser			
5	1940	1945	1950	
	CCC TAC GAC TCT CCT GAC CAG GAC CTG TTC TAT GCG ATC GCA GTT AAA	6084		
10	Pro Tyr Asp Ser Pro Asp Cln Asp Leu Phe Tyr Ala Ile Ala Val Lys			
	1955	1960	1965	
	GAT CTG ATA CGA AAG ACG GAC CCG AGC TAC AAA GTC AAG TCC CGC AAC	6132		
15	Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg Asn			
	1970	1975	1980	1985
	AGC ACC CTG GAG TAC ACC CTG AGC AAG CTG GAG CCC GGA GGG AAA TAC	6180		
20	Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys Tyr			
	1990	1995	2000	
	CAC GTC ATT GTG CAG CTG CGG AAC ATG AGC AAA GAT GCC AGT GTG AAG	6228		
25	His Val Ile Val Cln Leu Gly Asn Met Ser Lys Asp Ala Ser Val Lys			
	2005	2010	2015	
	ATC ACC ACC GTT TCG TTA TCG GCA CCC GAT GCC TTA AAA ATC ATA ACA	6276		
30	Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile Thr			
	2020	2025	2030	
	GAA AAT GAC CAC GTC CTT CTC TTC TGG AAA ACT CTA GCT CTA AAG GAA	6324		
35	Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys Glu			
	2035	2040	2045	
	AAG TAT TTT AAC GAA AGC AGG GGC TAC GAG ATA CAC ATG TTT GAT AGC	6372		
40	Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp Ser			
	2050	2055	2060	2065
	GCC ATG AAT ATC ACC GCA TAC CTT GGG AAT ACT ACT GAC AAT TTC TTT	6420		
45	Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe Phe			
	2070	2075	2080	
	AAA ATT TCC AAC CTG AAG ATG GGT CAC AAT TAC ACA TTC ACG GTC CAG	6468		
50	55			

Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val Gin
 5 2085 2090 2095
 GCA CGA TGC CTT TTG GGC AGC CAG ATC TGC GGG GAG CCT GCC GTG CTA 6516
 Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val Leu
 10 2100 2105 2110
 CTG TAT GAT GAG CTG GGG TCT GGT GGC GAT GCG TCG GCG ATG CAG GCT 6564
 Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gin Ala
 15 2115 2120 2125
 GCC AGG TCT ACT GAT CTC GCC CCC CTG GTG CCC ATC CTG TTT CTG 6612
 Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe Leu
 20 2130 2135 2140 2145
 ATA CTG CTG ACC CTG CGG GTC CGG TTT CCC ATC CTG TAC ACG AAG CAT 6660
 Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys His
 25 2150 2155 2160
 CCG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC ACC 6708
 Arg Arg Leu Gin Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr Ser
 30 2165 2170 2175
 TCC AGA CTC GGC TCC GCC ATC TTC TCC TCT CGG GAT GAC TTC GGG GAG 6756
 Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly Glu
 35 2180 2185 2190
 GAT GAT GAA GAT GCT CCT ATG ATC ACT GGA TTT TCG GAC GAC GTC CCC 6804
 Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val Pro
 40 2195 2200 2205
 ATG CTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA 6856
 Met Val Ile Ala
 45 2210
 TTTTATTTCA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT 6916
 GTTATTTTA TATGGGCCAA AAACAAAAGC AAAAAAAA AAAAA 6961
 50
 55

Sequence ID No. 4

Length of the Sequence: 300

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Sequence:

15 ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCGTGACC ATGGAGAGTG 60
ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTGTGGGC ATTTGACACC CACAAGCAAG 120
20 AGAGGAGAAC TTTGAACCTTC CGAGGAAGCA TATTGTCACA CAAACTTGGC AATCTGACAG 180
CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGAT AGCCCTCTGG 240
CATTTGAGCA TGTTATGACC AGAGGGGTTTC GCCCACCTGC ACCTAGCCTC AAGGCCAAAG 300

25 Sequence ID No. 5

Length of the Sequence: 6642

Type: nucleic acid

30 Strandedness: double

Topology: linear

35 Molecular type: cDNA to mRNA

Sequence:

40 ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCGT TCCTATTAC CCTGGTCGCA 60
CTGCTGCCGC CCGGAGCTCT CTGCGAAGTC TGGACCGAGA GGCTGCCACGG CGGCAGCGCG 120
CCCTTGCCCC AGGACCGGGG CTTCCCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG 180
TGGGCGCGCG GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CCGGAGGAAA 240
45 CGGAGCGCTG CCCTGCAGCC CGACCCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT 300
TOCCACAATC AGATGGTGCT GCACTGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG 360
50 GCGCGAGATA GCCTGGCATT CGCGAGGCCA AAGAGCAGTG ATGTGTACGT GTCTTACGAC 420
TATGGAAAAT CATTCAGAA AATTTAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT 480

55

	GAAGCTGTTA TCGGCCAGTT CTACCACAGC CCTGGGGACA ACAACGGTA CATCTTGC	540
5	GACGGTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTT	600
	TCCATCCCAT TTTCGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTG	660
	GGCTTTGACA GGTCCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC	720
10	TGGATCATGA TTCAGGAACA TGTCAACTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780
	CCAAATACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT	840
	ACAGATTTCT TCCAGTCCCC GGAAAACCAG GAACTGATCC TTGAGGAAGT GAGAGATTT	900
15	CAGCTTCGGG ACAAGTACAT GTTGTCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG	960
	CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAAGC CCATGAGAGC AGCCCAGTT	1020
20	CTCACAAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTCTT	1080
	GTGTGTGTCG CCCACAGTAA CAACCCGACC AATTATACCA TCTCAGAGGC AGAGGGGCTG	1140
	AACTTCTCCC TGTCTTGGG AACCGTGCTC TATTACAGCC CAGGAGGGGC CGGCAGTGAC	1200
25	ACCTTGGTGA CGTATTTGCA AAATGAACCA TTTGCTGACT TCCACCCAGT GGAAGGATTG	1260
	CAAGGAGTCT ACATTGCTAC TCTGATTAAT GGTTCTATGA ATGAGGAGAA CATGAGATCG	1320
	CTCATCACCT TTGACAAAGG CGGAACCTGG GAGTTTCTTC AGGCTCCAGC CTTCACCGGA	1380
30	TATGGAGAGA AAATCAATTG TCAAGCTTCC CAGGGCTGTT CCCTTCATCT GGCTCACCCG	1440
	CTCACTCAGC TCCTCAACCT CCAGCTCCGG AGAATGCCA TCCTGTCCAA GGAGTCGGCT	1500
	CCAGGCTCA TCATGCCAC TGGCTCAGTG GGAAAGAACT TGGCTAGCAA CACAAACGTG	1560
35	TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCTGGACC TCACTACTAC	1620
	ACATGGGGAG ACCACGGGGG AATCATCACG GCCATTGCC AGGGCATGGA ACCAACCGAG	1680
40	CTAAAATACA GTACCAATGA AGGGGAGACC TGGAAAACAT TCATCTCTC TGAGAACCCA	1740
	GTGTTTGTGT ATGGCTCCT CACAGAACCT GGGGAGAAGA GCACTGCTT CACCATCTT	1800
	GGCTCGAACCA AAGAGAAATGT CCACAGCTGG CTGATCCTCC AGGTCAATGC CACGGATCCC	1860
45	TTGGGAGTTC CCTGCACAGA GAATGACTAC AAGCTGTGGT CACCATCTGA TGAGGGGGGG	1920
	AATCACTGTT TGCTGGGACA CAAGACTGTT TTCAAACGGC GGACCCCCCA TCCCACATGC	1980
	TTCAATGGAG AGGACTTTGA CAGGCCGGTG GTCGTGTCCA ACTGCTCCTG CACCCGGGAG	2040
50	GACTATGAGT GTGACTTCGG TTTCAAGATG AGTGAAGATT TGTCAATTAGA CGTTTGTGTT	2100
	CCAGATCCGG AATTTCTGG AAAGTCATAC TCCCCTCCTG TCCCTGCCCG TGTGGGTTCT	2160

ACTTACAGGA GAACCGAGAGG CTACCGGAAG ATTTCTGGGG ACACMTGTAG CGGAGGAGAT 2220
 GTTGAAGCGC GACTCGAAGG AGAGCTGGTC CCCTGTCCCC TGGCAGAAGA GAACCGAGTC 2280
 ATTCTGTATG CTGTGAGGAA ATCCATCTAC CGCTATGACC TGGCCTCGGG AGCCACCCAG 2340
 CAGTTCCCTC TCACCCGGCT ACGGGCAGCA GTGGCCCTGG ACTTTGACTA TGAGCACAAC 2400
 10 TGTTTGTATT GGTCCGACCT GGCCCTGGAC GTCATCCAGC GCCTCTGTTT GAATGGAACC 2460
 ACAGGGCAAG AGGTGATCAT CAATTCTGGC CTGGAGACAG TAGAAGCTTT GGCTTTGAA 2520
 CCCCTCAGCC AGCTGCTTAA CTGGGTAGAT GCAGGGCTCA AAAAGATTGA CGTAGCTAAT 2580
 15 CCAGATGGCG ACTTCCGACT CACAATCGTC AATTCCCTCG TGCTTGATCG TCCCAGGGCT 2640
 CTGGTCCCTCG TCCCCAAGA CGGGGTGATG TTCTGGACAG ACTGGGGAGA CCTGAACCT 2700
 GGGATTTATC GGAGCAATAT GGATGGTTCT GCTGCCTATC ACCTGGTGTC TGAGGATGTG 2760
 20 AACTGGCCA ATGCCATCTC TGTGGACGAC CAGTGGATT ACTGGACGGA TGCCTACCTG 2820
 GAGTGCATAG AGCCGATCAC GTTCAGTGGC CAGCAGCGCT CTGTCATTCT GGACAACCTC 2880
 25 CGGCACCCCT ATGCCATTGC TGTCTTTAAG AATGAAATCT ACTGGGATGA CTGGTCACAG 2940
 CTCAGCATAT TCCGAGCTTC CAAATACAGT GGGTCCCAGA TCGAGATTCT GGCAAACCCAG 3000
 CTCACGGGGC TCATGGACAT GAAGATTTTC TACAAGGGGA AGAACACTGG AAGCAATGCC 3060
 30 TGTGTGCCCA GGCCATGCAG CCTGCTGTGC CTGGCCAAGG CCAACAAACAG TAGAAGCTGC 3120
 ACGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGGACCTGAT GTGTGACTGC 3180
 CCTCAGGGCT ATCACCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC 3240
 35 AACCAGTATC CCTCCAGCAA CGGGAACTGT ATCAACACCA TTTGGTGCTG TGACTTTGAC 3300
 AACGACTGTG GAGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC 3360
 40 ACCCAGTTTC GTTGGCAGGA GTCTGGACT TGTATCCCAC TGTCTATAA ATGTGACCTT 3420
 GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT 3480
 GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TCACGGGGAC 3540
 45 AACGACTGCA GGGACTGGTC TGATGAACCC AACTGTACCG CCATCTATCA CACCTGTGAG 3600
 GCCTCCAAC TCCAGTGCCG AAACGGGCAC TGCACTCCCC AGCCGTGGGC GTGTGACGGG 3660
 GATACTGGACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAACTGCAAT 3720
 50 CGATTCCGCT GCCCAAACGG CACTTGCATC CCATCCAGCA AACATGTGA TGGTCTCCGT 3780
 CATTGCTCTG ATGGCTCCGA TGAACAGCAC TGGGACCCCC TCTGTACGCCA CTTCATGGAC 3840

5 TTTGTGTGTA ACAACCCCA GCACTGCCCTG TTCCACTCCA TGGTCTGTCA CGGAATCATC 3900
 CAGTGGCGCG ACCGGTCCGA TGAGGATCCG GCGTTGCGAG CATGCTCCA AGATCCTGAG 3960
 TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGTAGA ATGGAGTGTG CATCAGTTG 4020
 ATTTGGAAGT GCGACGGGAT GGATGATTGC GCGGATTATT CTGATGAACC CAACTGCGAA 4080
 10 AACCCCACAG AAGCCCCAAA CTGCTCCCAG TACTTCCAGT TTCCGTGTGA GAATGCCAC 4140
 TGCATCCCCA ACAGATGAA ATGTGACAGG GAGAACGACT GTGGGACTG STCTGATGAG 4200
 AAGGATTGTG GAGATTACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG 4260
 15 CCCAATTACT ACCGCTGCAG CACTGGGACC TCCGTGATGG ACACCTGGCT GTGGACGGG 4320
 TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCTTGC AACAGTCACT 4380
 GCTGCCTCCA CTCCCACCC AACTGGGCA TGTGACCGAT TTGAGTTCCA ATGCCACCAA 4440
 20 CGGAAGACGT GTATTCCAA CTGGAAGGCC TGTGACGGCC ACCAAGATG CCAGGATGGC 4500
 CGGGACGAGG CCAATTGCC CACACACAGC ACCTTGACTT CCATGAGCAG GGAGTTCCAG 4560
 25 TGCGAGGACG GGGAGGCCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC 4620
 TCGGACGAGA GCGATGAAAAA GGCTGCACT GATGAGTCA CTGTGTACAA ACTACAGAAT 4680
 CTTCACTGGA CAGCTGACTT CTCTGGGAT GTGACTTGA CCTGGATGAG CCCAAAAAA 4740
 30 ATGCCCTCTG CATCTGTGT ATATAATGTC TACTACAGG TGGTGGAGA GACCATATGG 4800
 AAGACTCTGG AGACCCACAG CAATAAGACA AACACTSTAT TAAAAGTCTT GAAACCAGAT 4860
 ACCACGTATC AGGTTAAAGT ACAGGTTCACT TGTCTCACCA AGGCACACAA CACCAATGAC 4920
 35 TTTGTGACCC TGAGGACCCC AGAGGATTC CCAGATCCC CTCGAAATCT CCAGCTGTCA 4980
 CTCCCCAGGG AAGCAGAAGG TGTGATTGTA GGCCACTGGG CTCCTCCCAT CCACACCCAT 5040
 40 GGCCTCATCC CTGAGTACAT TGTAGAATAC AGCAGGAGTG GTTCCAAGAT GTGGGCTCC 5100
 CAGAGGGCTG CTAGTAACCT TACAGAAATC AAGAACTTAT TGGTCAACAC TCTATAACCC 5160
 GTCAGACTGG CTGGGTGAC TACTCGTGA ATAGGAAACT GGACCGATTC TAAATCCATT 5220
 45 ACCACCATAA AAGGAAAAGT GATCCCACCA CCAGATATCC ACATTGACAG CTATGGTCAA 5280
 AATTATCTAA GCTTCACCT GACCATGGAG AGTGTATCA AGGTGAATGG CTATGTGGTG 5340
 AACCTTTCTT GGGCATTGAA CACCCACAAG CAAGAGAGGA GAACTTGAA CTTCCGAGGA 5400
 50 AGCATATTGT CACACAAAGT TGGCAATCTG ACAGCTCATA CATCCTATGA GATTCTGCC 5460
 TGGGCAAGA CTGACTTGGG CGATAGCCCT CTGGCATTTG ACCATGTTAT GACCAGAGGG 5520

55

5 GTTCGGCCAC CTGCACCTAG CCTCAAGGCC AAAGCCATCA ACCAGACTGC ACTGGAATCT 5580
 CCTGGACCG GCCCCCGGAA TGCGGTTAT GGTATTTCT ATGCCACGTC CTMTCMTGAC 5640
 CTCTATCCA ACCCGAAGAG CTTGACTACT TCACTCCACA ACAAGACGGT CATTGTCAGT 5700
 10 AAGGATGAGC AGTATTTGTT TCTGGTCCGT GTACTGGTAC CCTACCAGGG GCCATCCTCT 5760
 GACTACGTTG TAGTGAAGAT GATCCCCGAC ACCAGGCTTC CACCCCGTCA CCTGCATGTG 5820
 15 GTTCATACGG GCAAAACCTC CGTGGTCATC AAGTGGAAAT CACCGTATGA CTCTCCGAC 5880
 CAGGACTTGT TGTATGCAAT TGCAGTCAAA GATCTCATAA GAAAGACTGA CAGGAGCTAC 5940
 20 AAAGTAAAAT CCCGTAACAG CACTGTGGAA TACACCCCTTA ACAAGTTGGA GCCTGGGGGG 6000
 AAATACCACA TCATTGCTCA ACTCGGGAAC ATGAGCAAAG ATTCCAGCAT AAAAATTACC 6060
 25 ACAGTTTCAT TATCACCACC TGATGCCCTTA AAAATCATAA CAGAAAATGA TCATGTTCTT 6120
 CTGTTTTGGA AAAGCCTGGC TTAAAGGAA AACCAATTATA ATGAAAGCAG GGGCTATGAG 6180
 ATACACATCT TTGATACTGC CATGAATATC ACAGCTTACC TTGGGAATAC TACTGACAAT 6240
 30 TTCTTTAAAA TTTCCAACCT GAAGATGGCT CATAATTACA CGTTCACCGT CCAAGCAAGA 6300
 TGCCCTTTTG GCAACCAGAT CTGCGGGAG CCTGCCATCC TGCTGTACGA TGAGCTGGGG 6360
 TCTCGTCCAG ATGCATCTGC AACCGAGGCT CCCAGATCTA CGGATGTTGC TGCTGTGGTG 6420
 35 GTGCCCATCT TATTCCGTAT ACTGCTGAGC CTGGGGGTGG GGTTTGCCAT CCTGTACAG 6480
 AAGCACCGGA GGCTGCAGAG CAGCTTCACC GCCTTCGCCA ACAGCCACTA CAGCTCCAGG 6540
 CTGGGGTCCG CAATCTTCTC CTCTGGGAT GACCTGGGGG AACATGATGA AGATGCCCT 6600
 40 ATGATAACTG GATTTTCAGA TGACGTCCCC ATGGTGATAG CC 6642

Sequence ID No. 6

Length of the Sequence: 2214

Type: amino acid

Topology: linear

Molecular type: Protein

Sequence:

Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
 50 5 10 15
 Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr

55

	20	25	30
5	Gln Arg Leu His Gly Cys Ser Ala Pro Leu Pro Gln Asp Arg Glu Phe		
	35	40	45
	Leu Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly		
10	50	55	60
	Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys		
	65	70	75
15	Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val		
	85	90	95
20	Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu		
	100	105	110
	Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala		
25	115	120	125
	Arg Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser		
	130	135	140
30	Phe Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser		
	145	150	155
	Glu Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg		
35	165	170	175
	Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp		
40	180	185	190
	Phe Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp		
	195	200	205
45	Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg		
	210	215	220
	Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr		
50	225	230	235
	Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp		

55

	245	250	255
5	Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser		
	260	265	270
	Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu		
10	275	280	285
	Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp		
	290	295	300
15	Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln		
	305	310	315
	Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg		
20	325	330	335
	Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala		
25	340	345	350
	Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn		
	355	360	365
30	Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu		
	370	375	380
	Ser Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp		
35	385	390	395
	Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg		
	405	410	415
40	Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser		
	420	425	430
45	Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly		
	435	440	445
	Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys		
50	450	455	460
	Ile Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg		
55			

465	470	475	480	
Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser				
5	485	490	495	
Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys				
10	500	505	510	
Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ala Gly Ala				
15	515	520	525	
Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Thr Trp Gly Asp				
20	530	535	540	
His Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu				
25	545	550	555	560
Leu Lys Tyr Ser Thr Asn Glu Gly Gln Thr Trp Lys Thr Phe Ile Phe				
30	565	570	575	
Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu				
35	580	585	590	
Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His				
40	595	600	605	
Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro				
45	610	615	620	
Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly				
50	625	630	635	640
Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro				
55	645	650	655	
His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val				
60	660	665	670	
Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe				
65	675	680	685	
Lys Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu				

690	695	700
Phe Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser		
705	710	715
Thr Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys		
725	730	735
Ser Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys		
740	745	750
Pro Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser		
755	760	765
Ile Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu		
770	775	780
Thr Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn		
785	790	795
Cys Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys		
805	810	815
Leu Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu		
820	825	830
Thr Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp		
835	840	845
Val Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp		
850	855	860
Phe Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala		
865	870	875
Leu Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly		
885	890	895
Asp Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala		
900	905	910
Tyr His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val		

55

EP 0 773 290 A2

915 920 925
5 Asp Asp Gln Trp Ile Tyr Trp Thr Asp Ala Ty: Leu Glu Cys Ile Glu
930 935 940
Arg Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu
10 945 950 955 960
Pro His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp
965 970 975
15 Asp Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser
980 985 990
Gln Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys
20 995 1000 1005
Ile Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg
1010 1015 1020
25 Pro Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys
1025 1030 1035 1040
30 Arg Cys Pro Glu Asp Val Ser Ser Val Leu Pro Ser Gly Asp Leu
1045 1050 1055
35 Met Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val
1060 1065 1070
Lys Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly
1075 1080 1085
40 Asn Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly
1090 1095 1100
45 Asp Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp
1105 1110 1115 1120
50 Thr Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr
1125 1130 1135
Lys Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His

55

	1140	1145	1150	
5	Cys Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly			
	1155	1160	1165	
	Met Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg			
10	1170	1175	1180	
	Asp Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu			
	1185	1190	1195	1200
15	Ala Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp			
	1205	1210	1215	
	Ala Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro			
20	1220	1225	1230	
	Val Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr			
	1235	1240	1245	
25	Cys Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp			
	1250	1255	1260	
30	Gly Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp			
	1265	1270	1275	1280
	Phe Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys			
35	1285	1290	1295	
	Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe			
	1300	1305	1310	
40	Ala Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe			
	1315	1320	1325	
	Gly Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys			
45	1330	1335	1340	
	Asp Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu			
50	1345	1350	1355	1360
	Asn Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys			

55

	1365	1370	1375
5	Glu Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn		
	1380	1385	1390
10	Asp Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile		
	1395	1400	1405
15	Leu Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr		
	1410	1415	1420
20	Arg Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly		
	1425	1430	1435
25	Tyr Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu		
	1440	1445	1450
30	Ala Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp		
	1455	1460	1465
35	Arg Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp		
	1470	1475	1480
40	Lys Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala		
	1485	1490	1495
45	Asn Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln		
	1500	1505	1510
50	1515		
	1520	1525	1530
55	Cys Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly		
	1535	1540	1545
60	Phe Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu		
	1550	1555	1560
65	Leu Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser		
	1565	1570	1575
70	Gly Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala		
	1580	1585	1590
75	Ser Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp		

1585	1590	1595	1600
Lys Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val			
5	1605	1610	1615
Leu Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu			
10	1620	1625	1630
Ser Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu			
15	1635	1640	1645
Gly Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu			
20	1650	1655	1660
Ala Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His			
25	1665	1670	1675
Gly Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys			
30	1685	1690	1695
Met Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn			
35	1700	1705	1710
Leu Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser			
40	1715	1720	1725
Arg Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys			
45	1730	1735	1740
Gly Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu			
50	1745	1750	1755
Asn Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn			
55	1765	1770	1775
Gly Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Gln Glu			
60	1780	1785	1790
Arg Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly			
65	1795	1800	1805
Asn Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr			

1810	1815	1820
Asp Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly		
1825	1830	1835
Val Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr		
1840	1845	1850
Ala Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile		
1855	1860	1865
Phe Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu		
1870	1875	1880
Thr Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Cln		
1885	1890	1895
Tyr Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser		
1900	1905	1910
1915		
1920		
Asp Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg		
1925	1930	1935
His Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp		
1940	1945	1950
Glu Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala		
1955	1960	1965
Val Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser		
1970	1975	1980
Arg Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly		
1985	1990	1995
2000		
Lys Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser		
2005	2010	2015
Ile Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile		
2020	2025	2030
Ile Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu		

	2035	2040	2045
5	Lys Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe		
	2050	2055	2060
10	Asp Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn		
	2065	2070	2075
15	Phe Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr		
	2085	2090	2095
20	Val Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala		
	2100	2105	2110
25	Ile Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr		
	2115	2120	2125
30	Gln Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu		
	2130	2135	2140
35	Phe Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr		
	2145	2150	2155
40	Lys His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His		
	2165	2170	2175
45	Tyr Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu		
	2180	2185	2190
50	Gly Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp		
	2195	2200	2205
55	Val Pro Met Val Ile Ala		

Sequence ID No. 7

Length of the Sequence: 6843

Type: nucleic acid

Strandedness: double

Topology: linear

Molecular type: cDNA to mRNA

Feature:

5 Name/Key: sig peptide

Location: 81..164

10 Identification method: S

Name/Key: mat peptide

Location: 165..6722

15 Identification method: S

Sequence:

20 CCG GCCCAGCGGC TCTCCTGGCC 23

TCGGCGCTCCA CATTCTCTCC TGGCGGGGGC CCCACCTGCA GTAGCGTTCG CCCGAACATG 33

Met

1

25 GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCG TTC CTA TTC ACC 131

Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr

30 5 10 15

CTG CTG GCA CTG CTG CCG CCC GGA CCT CTC TGC GAA GTC TGG ACG CAG 179

Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Gln

35 20 25 30

40 AGG CTG CAC GGC GGC AGC GCG CCC TTG CCC CAG GAC CCG GGC TTC CTC 227

Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe Leu

45 35 40 45

45 GTG CTG CAC GGC GAC CCG CGC GAG CTG CGG CTC TCG CGG CGC GCG GAT 275

Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly Asp

50 50 55 60 65

55 GGC AGG GGG GCG AGC CGC GCG GAC GAG AAG CCG CTC CGG AGG AAA CGG 323

50 Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys Arg

70 75 80

55

5	ACC GCT GCC CTG CAG CCC GAG CCC ATC AAG CTG TAC GGA CAG GTT AGT	371
	Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser	
	85 90 95	
10	CTG AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCT GGA GAG AAA	419
	Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys	
	100 105 110	
15	AGC AAC GTG ATC GTG GCC TTG GCC CGA GAT AGC CTG GCA TTG GCG AGG	467
	Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg	
	115 120 125	
20	CCC AAG AGC AGT GAT GTG TAC GTG TCT TAC GAC TAT GGA AAA TCA TTC	515
	Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe	
	130 135 140 145	
25	AAG AAA ATT TCA GAC AAG TTA AAC TTT GGC TTG GGA AAT AGG AGT GAA	563
	Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser Glu	
	150 155 160	
30	GCT GTT ATC CCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAG CGG TAC	611
	Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr	
	165 170 175	
35	ATC TTT GCA GAC GCT TAT GCC CAG TAC CTC TGG ATC ACG TTT GAC TTC	659
	Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe	
	180 185 190	
40	TGC AAC ACT CTT CAA CGC TTT TCC ATC CCA TTT CGG GCA CCT GAT CTC	707
	Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu	
	195 200 205	
45	CTC CTA CAC ACT AAG GCC TCC AAC CTT CTC TTG GGC TTT GAC AGG TCC	755
	Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser	
	210 215 220 225	
50	CAC CCC AAC AAG CAG CTG TGG AAG TCA GAT GAC TTT GGC CAG ACC TGG	803

	His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp			
5	230	235	240	
	ATC ATG ATT CAG GAA CAT GTC AAG TCC TTT TCT TGG GGA ATT GAT CCC		851	
	Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro			
10	245	250	255	
	TAT GAC AAA CCA AAT ACC ATC TAC ATT GAA CGA CAC GAA CCC TCT CGC		899	
	Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly			
15	260	265	270	
	TAC TCC ACT GTC TTC CGA ACT ACA GAT TTC TTC CAG TCC CGG GAA AAC		947	
	Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn			
20	275	280	285	
	CAG GAA GTG ATC CTT GAG GAA GTG AGA GAT TTT CAG CTT CGG GAC AAG		995	
	Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys			
25	290	295	300	305
	TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG CGC ACT GAA CAG CAG		1043	
	Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln Gln			
30	310	315	320	
	TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG ACA GCA		1091	
	Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala			
35	325	330	335	
	GCC CAG TTT GTC ACA AGA CAT CCT ATT AAT GAA TAT TAC ATC GCA GAT		1139	
	Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp			
40	340	345	350	
	GCC TCC GAG GAC CAG GTG TTT GTG TGT STC ACC CAC ACT AAC AAC CGC		1187	
	Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg			
45	355	360	365	
	ACC AAT TTA TAC ATC TCA GAG GCA GAG GGG CTG AAG TTC TCC CTG TCC		1235	
	Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser			
50				
55				

370	375	380	385	
TTG GAG AAC GTG CTC TAT TAC AGC CCA GGA GGG GCC GGC AGT GAC ACC				1283
Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp Thr				
390	395	400		
TTC CTG AGG TAT TTT GCA AAT GAA CCA TTT GCT GAC TTC CAC CGA CTG				1331
Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val				
405	410	415		
GAA GGA TTG CAA GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG				1379
Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met				
420	425	430		
AAT GAG GAG AAC ATG AGA TCG GTC ATC ACC TTT GAC AAA GGG GGA ACC				1427
Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Thr				
435	440	445		
TCG GAG TTT CTT CAG CCT CCA GCC TTC ACG GGA TAT GGA GAG AAA ATC				1475
Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile				
450	455	460	465	
AAT TGT GAG CTT TCC CAG GGC TGT TCC CTT CAT CTG GCT CAG CGC CTC				1523
Asn Cys Glu Leu Ser Gln Cys Ser Leu His Leu Ala Gln Arg Leu				
470	475	480		
AGT CAG CTC CTC AAC CTC CAG CTC CGG AGA ATG CCC ATC CTG TCC AAG				1571
Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys				
485	490	495		
CAG TCG GCT CCA GGC CTC ATC ATC GCC ACT GGC TCA GTG GGA AAG AAC				1619
Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn				
500	505	510		
TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG				1667
Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg				
515	520	525		

55

	TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAC ACA TGG GGA GAC CAC		1715	
5	Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His			
	530	535	540	545
	GCC CGA ATC ATC ACG GCC ATT GCC CAG GGC ATG GAA ACC AAC GAG CTA		1763	
10	Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu			
	550	555	560	
	AAA TAC ACT ACC AAT GAA GGG GAG ACC TGG AAA ACA TTC ATC TTC TCT		1811	
15	Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe Ser			
	565	570	575	
20	GAG AAG CCA GTG TTT GTG TAT GCC CTC CTC ACA GAA CCT CGG GAG AAC		1859	
	Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys			
	580	585	590	
25	AGC ACT GTC TTC ACC ATC TTT CGC TCG AAC AAA GAG AAT GTC CAC AGC		1907	
	Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser			
	595	600	605	
30	TGG CTC ATC CTC CAG GTC AAT GCC ACC GAT GCC TTG GGA GTT CCC TGC		1955	
	Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys			
	610	615	620	625
35	ACA GAG AAT GAC TAC AAG CTG TGG TCA CCA TCT GAT GAC CGG GGG AAT		2003	
	Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn			
	630	635	640	
40	GAG TGT TTG CTG GGA CAC AAG ACT GTT TTC AAA CGG CGG ACC CCC CAT		2051	
	Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His			
	645	650	655	
45	CCC ACA TCC TTC AAT GGA GAG GAC TTT GAC AGG CGG GTG GTC GTG TCC		2099	
	Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser			
50	660	665	670	
	AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTC GGT TTC AAG		2147	

Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Lys
 675 680 685
 5 ATG AGT GAA GAT TTG TCA TTA GAG GTT TGT GTT CCA GAT CCG GAA TTT 2195
 Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu Phe
 10 690 695 700 705
 TCT GGA AAG TCA TAC TCC CCT CCT GTG CCT TGC CCT GTG GGT TCT ACT 2243
 Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
 15 710 715 720
 TAC AGG AGA ACC AGA GCC TAC CCG AAG ATT TCT CGG GAC ACT TGT ACC 2291
 Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
 20 725 730 735
 CGA CGA GAT GTT GAA CGG CGA CTG CAA CGA GAG CTC GTC CCC TGT CCC 2339
 Gly Gly Asp Val Glu Ala Arg Leu Glu Leu Val Pro Cys Pro
 25 740 745 750
 CTG CCA GAA GAG AAC GAG TTC ATT CTG TAT GCT GTG AGG AAA TCC ATC 2387
 Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser Ile
 30 755 760 765
 TAC CGC TAT GAC CTG GCC TCG CGA GCC ACC GAG CAG TTG CCT CTC ACC 2435
 Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu Thr
 35 770 775 780 785
 CGG CTA CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGT 2483
 Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
 40 790 795 800
 TTG TAT TGG TCC GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG 2531
 Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
 45 805 810 815
 AAT CGA AGC ACA GGG CAA GAG GTG ATC ATC AAT TCT GGC CTG GAG ACA 2579
 Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu Thr
 50
 55

EP 0 773 290 A2

	820	825	830	
5	GTA GAA GCT TTG GCT TTT GAA CCC CTC AGC CAG CTG CTM TAC TCG GTA			2627
	Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val			
	835	840	845	
10	GAT GCA GGC TTC AAA AAG ATT GAG GTA GCT AAT CCA GAT GGC GAC TTC			2675
	Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe			
	850	855	860	865
15	CGA CTC ACA ATC GTC AAT TCC TCT GTG CTT GAT CGT CCC AGG GCT CTG			2723
	Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu			
	870	875	880	
20	GTC CTC GTG CCC CAA GAG GGG GTC ATG TTC TGG ACA GAC TGG GGA GAC			2771
	Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly Asp			
	885	890	895	
25	CTG AAG CCT CGG ATT TAT CCG AGC AAT ATG GAT CGT TCT GCT GCC TAT			2819
	Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr			
	900	905	910	
30	CAC CTG GTG TCT GAG GAT GTG AAG TGG CCC AAT CGC ATC TCT GTG GAC			2867
	His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp			
	915	920	925	
35	GAC CAG TGG ATT TAC TGG ACG GAT GCC TAC CTG GAG TGC ATA GAG CGG			2915
	Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu Arg			
	930	935	940	945
40	ATC ACG TTC AGT GGC CAG CAG CGC TCT GTC ATT CTG GAC AAC CTC CGG			2963
	Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu Pro			
	950	955	960	
45	CAC CCC TAT GCC ATT GCT GTC TTT AAG AAT GAA ATC TAC TGG GAT GAC			3011
	His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp			
	965	970	975	

	TGG TCA CAG CTC AGC ATA TTC CGA GCT TCC AAA TAC AGT GGG TCC CAG	3059
5	Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln	
	980 985 990	
	ATG GAG ATT CTG GCA AAC CAG CTC ACG GGG CTC ATG GAC ATG AAG ATT	3107
10	Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys Ile	
	995 1000 1005	
	TTC TAC AAG GGG AAG AAC ACT GGA AGC AAT GCC TGT GTG CCC AGG CCA	3155
15	Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro	
	1010 1015 1020 1025	
	TGC AGC CTG CTG TGC CTG CCC AAG GCC AAC AAC ACT AGA AGC TGC AGG	3203
20	Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys Arg	
	1030 1035 1040	
	TGT CCA GAG GAT GTG TCC AGC AGT GTG CTT CCA TCA GGG GAC CTG ATG	3251
25	Cys Pro Glu Asp Val Ser Ser Val Leu Pro Ser Gly Asp Leu Met	
	1045 1050 1055	
	TGT GAC TGC CCT CAG GGC TAT CAG CTC AAG AAC AAT ACC TGT GTC AAA	3299
30	Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val Lys	
	1060 1065 1070	
	CAA GAG AAC ACC TGT CTT CGC AAC CAG TAT CGC TGC AGC AAC GGG AAC	3347
35	Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn	
	1075 1080 1085	
	TGT ATC AAC AGC ATT TGG TGG TGT GAC TTT GAC AAC GAC TGT GGA GAC	3395
40	Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp	
	1090 1095 1100 1105	
	ATG AGC GAT GAG AGA AAC TGC CCT ACC ACC ATC TGT GAC CTG GAC ACC	3443
45	Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr	
	1110 1115 1120	
50	CAG TTT CGT TGC CAG GAG TCT GGG ACT TGT ATC CCA CTG TCC TAT AAA	3491

55

Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys
 1125 1130 1135
 5 TGT GAC CTT GAG GAT GAC TGT GGA GAC AAC ACT GAT GAA AGT CAT TGT 3539
 Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His Cys
 1140 1145 1150
 10 GAA ATG CAC CAG TGC CGG AGT GAC GAG TAC AAC TGC AGT TCC GGC ATG 3587
 Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met
 1155 1160 1165
 15 TGC ATC CGC TCC TCC TGG GTC TGT GAC GGG GAC AAC GAC TGC AGG GAC 3635
 Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp
 1170 1175 1180 1185
 20 TCG TCT GAT GAA GCC AAC TGT ACC GCC ATC TAT CAC ACC TGT GAG GCC 3683
 Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala
 1190 1195 1200
 25 TCC AAC TTC CAG TGC CGA AAC GGG CAC TGC ATC CCC CAG CGG TCG GCG 3731
 Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala
 1205 1210 1215
 30 TGT GAC GGG GAT ACG GAC TGC CAG CAT CGT TCC GAT CAG GAT CCA GTC 3779
 Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Val
 1220 1225 1230
 35 AAC TCT GAG AAC AAG TGC AAT GGA TTC CGC TCC CCA AAC GGC ACT TGC 3827
 Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys
 1235 1240 1245
 40 ATC CCA TCC AGC AAA CAT TGT GAT CGT CTG CCT GAT TGC TCT GAT GGC 3875
 Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp Gly
 1250 1255 1260 1265
 45 TCC GAT GAA CAG CAC TCC GAG CCC CTC TGT ACG CAC TTC ATG GAC TTT 3923
 Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp Phe
 50

	1270	1275	1280	
5	CTG TGT AAG AAC CGC CAG CAG TGC CTG TTC CAC TCC ATG GTC TGT GAC			3971
	Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp			
10	1285	1290	1295	
	GGA ATC ATC CAG TGC CGC GAC GGG TCC GAT GAG GAT GCG GCG TTT GCA			4019
	Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe Ala			
15	1300	1305	1310	
	GGA TGC TCC CAA GAT CCT GAG TTC CAC AAG GTC TGT GAT GAG TTC GGT			4067
	Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly			
20	1315	1320	1325	
	TTC CAG TGT CAG AAT GGA GTG TGC ATC AGT TTG ATT TGG AAG TGC GAC			4115
	Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp			
25	1330	1335	1340	1345
	GGG ATG GAT GAT TGC GGC GAT TAT TCT GAT GAA GCC AAC TGC GAA AAC			4163
	Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn			
30	1350	1355	1360	
	CCC ACA GAA GCC CCA AAC TGC TCC CGC TAC TTC CAG TTT CGG TGT GAG			4211
	Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Glu			
35	1365	1370	1375	
	AAT GGC CAC TGC ATC CCC AAC AGA TGG AAA TGT GAC AGG GAG AAC GAC			4259
	Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp			
40	1380	1385	1390	
	TGT GGG GAC TGG TCT GAT GAG AAG GAT TCT GGA GAT TCA CAT ATT CTT			4307
	Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile Leu			
45	1395	1400	1405	
	CCC TTC TCG ACT CCT GGG CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC			4355
	Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg			
50	1410	1415	1420	1425

55

5	TGC AGC AGT GGG ACC TGC GTG ATG GAC ACC TGG GTC TGC GAC GGG TAC	4403
	Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly Tyr	
	1430 1435 1440	
10	CCA GAT TGT GCA GAT GGC TCT GAC GAG GAA GCC TGC CCC TTG CTT GCA	4451
	Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu Ala	
	1445 1450 1455	
15	AAC GTC ACT GCT GCC TCC ACT CCC ACC CAA CTT GGG CGA TGT GAC CGA	4499
	Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp Arg	
	1460 1465 1470	
20	TTT GAG TTC GAA TGC CAC CAA CCG AAG ACG TGT ATT CCC AAC TGG AAG	4547
	Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp Lys	
	1475 1480 1485	
25	CGC TCT GAC GGC CAC CAA GAT TGC CAG GAT GCC CGG GAC GAG GCC AAT	4595
	Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala Asn	
	1490 1495 1500 1505	
30	TGC CCC ACA CAC AGC ACC TTG ACT TGC ATG ACC AGG GAG TTC CAG TGC	4643
	Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln Cys	
	1510 1515 1520	
35	GAG GAC GGG CAG GCC TGC ATT GTG CTC TCG GAG CGC TGC GAC GGC TTC	4691
	Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe	
	1525 1530 1535	
40	CTG GAC TGC TCG GAC GAG AGC GAT GAA AAG GCC TGC AGT GAT GAG TTG	4739
	Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu	
	1540 1545 1550	
45	ACT CTG TAC AAA GTC CAG AAT CTT CAG TCG ACA CCT GAC TTC TCT CCC	4787
	Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly	
	1555 1560 1565	
50	GAT GTG ACT TTG ACC TGG ATG AGG CCC AAA AAA ATG CCC TCT GCA TCT	4835

55

Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ser
 1570 1575 1580 1585
 TGT GTA TAT AAT GTC TAC TAC AGG GTG GTT GGA GAG AGC ATA TGG AAG 4883
 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
 1590 1595 1600
 ACT CTG GAG ACC CAC AGC AAT AAC ACA AAC ACT GTA TTA AAA GTC TTG 4931
 Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
 1605 1610 1615
 AAA CCA GAT ACC ACG TAT CAG GTT AAA GTC CAG GTT CAG TGT CTC ACC 4979
 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
 1620 1625 1630
 AAG GCA CAC AAC ACC AAT GAC TTT GTG ACC CTG AGG ACC CCA GAG GCA 5027
 Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
 1635 1640 1645
 TTG CCA GAT GCC CCT CGA AAT CTC CAG CTG TCA CTC CCC AGG GAA GCA 5075
 Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu Ala
 1650 1655 1660 1665
 CAA GGT GTG ATT GTC GGC CAC TGG GCT CCT CCC ATC CAC ACC CAT GGC 5123
 Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His Gly
 1670 1675 1680
 CTC ATC CGT GAG TAC ATT GTC GAA TAC AGC AGG AGT GGT TCC AAG ATG 5171
 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Met
 1685 1690 1695
 TGG GCC TCC CAG AGG GCT GCT AGT AAC TTT ACA GAA ATC AAG AAC TTA 5219
 Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
 1700 1705 1710
 TTG GTC AAC ACT CTA TAC ACC GTC AGA GTG GCT GCG GTG ACT ACT CGT 5267
 Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg

55

	1715	1720	1725	
5	GGA ATA GGA AAC TGG AGC GAT TCT AAA TCC ATT ACC ACC ATA AAA GGA			5315
	Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys Gly			
	1730	1735	1740	1745
10	AAA GTG ATC CCA CCA GAT ATC CAC ATT GAC AGC TAT GGT GAA AAT			5363
	Lys Val Ile Pro Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu Asn			
	1750	1755	1760	
15	TAT CTA AGC TTC ACC CTG ACC ATG GAG AGT GAT ATC AAC GTG AAT GGC			5411
	Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn Gly			
	1765	1770	1775	
20	TAT GTG GTG AAC CTT TTC TCG GCA TTT GAC ACC CAC AAG CAA GAG AGG			5459
	Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Arg			
	1780	1785	1790	
25	AGA ACT TTG AAC TTC CGA GGA AGC ATA TTG TCA CAC AAA GTT GGC AAT			5507
	Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly Asn			
	1795	1800	1805	
30	CTG ACA GCT CAT ACA TCC TAT GAG ATT TCT GCC TGG GCC AAG ACT GAC			5555
	Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp			
	1810	1815	1820	1825
35	TTG GGG GAT AGC CCT CTG GCA TTT GAG CAT GTT ATG ACC AGA GGG GTT			5603
	Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly Val			
	1830	1835	1840	
40	CGC CCA CCT GCA CCT AGC CTC AAG GCC AAA GCC ATC AAC CAG ACT GCA			5651
	Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr Ala			
	1845	1850	1855	
45	GTG GAA TGT ACC TGG ACC GGC CCC CGG AAT GTG GTT TAT GGT ATT TTC			5699
	Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe			
	1860	1865	1870	

5	TAT	GCC	ACG	TCC	TTT	CTT	GAC	CTC	TAT	CGC	AAC	CCG	AAG	AGC	TTG	ACT	5747
	Tyr	Ala	Thr	Ser	Phe	Leu	Asp	Leu	Tyr	Arg	Asn	Pro	Lys	Ser	Leu	Thr	
10	1875				1880				1885								5795
	ACT	TCA	CTC	CAC	AAC	AAG	ACG	GTC	ATT	GTC	AGT	AAG	GAT	GAG	CAG	TAT	
15	Thr	Ser	Leu	His	Asn	Lys	Thr	Val	Ile	Val	Ser	Lys	Asp	Glu	Gln	Tyr	
20	1890			1895				1900			1905						5843
	TTG	TTT	CTG	GTC	CGT	GTA	GTG	GTA	CCC	TAC	CAG	GGG	CCA	TCC	TCT	GAC	
25	Leu	Phe	Leu	Val	Arg	Val	Val	Val	Pro	Tyr	Gln	Gly	Pro	Ser	Ser	Asp	
30	1910			1915			1920										5891
	TAC	CTT	GTA	GTG	AAG	ATG	ATC	CCG	GAC	AGC	AGG	CTT	CCA	CCC	CGT	CAC	
35	Tyr	Val	Val	Val	Lys	Met	Ile	Pro	Asp	Ser	Arg	Leu	Pro	Pro	Arg	His	
40	1925			1930			1935										5939
	CTG	CAT	CTG	GTT	CAT	ACG	GGC	AAA	ACC	TCC	GTG	GTC	ATC	AAG	TGG	GAA	
45	Leu	His	Val	Val	His	Thr	Gly	Lys	Thr	Ser	Val	Val	Ile	Lys	Trp	Glu	
50	1940			1945			1950										5987
	TCA	CCG	TAT	GAC	TCT	CCT	CAC	CAG	SAC	TTG	TTG	TAT	GCA	ATT	GCA	GTC	
55	Ser	Pro	Tyr	Asp	Ser	Pro	Asp	Gln	Asp	Leu	Leu	Tyr	Ala	Ile	Ala	Val	
60	1955			1960			1965										6035
	AAA	GAT	CTC	ATA	AGA	AAG	ACT	GAC	AGC	TAC	AAA	GTA	AAA	TCC	CGT		
65	Lys	Asp	Leu	Ile	Arg	Lys	Thr	Asp	Arg	Ser	Tyr	Lys	Val	Lys	Ser	Arg	
70	1970			1975			1980			1985							6083
	AAC	AGC	ACT	GTG	GAA	TAC	ACC	CTT	AAC	AAG	TTG	GAG	CCT	GGC	GGG	AAA	
75	Asn	Ser	Thr	Val	Glu	Tyr	Thr	Leu	Asn	Lys	Leu	Glu	Pro	Gly	Gly	Lys	
80	1990			1995			2000										6131
	TAC	CAC	ATC	ATT	GTC	CAA	CTG	GGG	AAC	ATG	AGC	AAA	GAT	TCC	AGC	ATA	
85	Tyr	His	Ile	Ile	Val	Gln	Leu	Gly	Asn	Met	Ser	Lys	Asp	Ser	Ser	Ile	
90	2005			2010			2015										6179
	AAA	ATT	ACC	ACA	GTT	TCA	TTA	TCA	GCA	CCT	GAT	GCC	TTA	AAA	ATC	ATA	
95																	

	Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile			
5	2020	2025	2030	
	ACA GAA AAT GAT CAT GTT CTT CTG TTT TGG AAA AGC CTG GCT TTA AAG		6227	
	Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys			
10	2035	2040	2045	
	GAA AAG CAT TTT AAT GAA AGC AGG GGC TAT GAG ATA CAC ATG TTT GAT		6275	
	Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp			
15	2050	2055	2060	2065
	ACT GCC ATG AAT ATC ACA GCT TAC CTT GGG AAT ACT ACT GAC AAT TTC		6323	
20	Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe			
	2070	2075	2080	
	TTT AAA ATT TCC AAC CTG AAG ATG GGT CAT AAT TAC ACG TTC ACC GTC		6371	
25	Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val			
	2085	2090	2095	
	CAA GCA AGA TGC CTT TTT GGC AAC CAG ATC TGT GGG GAG CCT GCC ATC		6419	
30	Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala Ile			
	2100	2105	2110	
	CTG CTG TAC GAT GAG CTG GGG TCT GGT GCA GAT CCA TCT GCA ACG CAG		6467	
35	Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr Gln			
	2115	2120	2125	
	GCT GCC AGA TCT ACG GAT GTT GCT GCT GTG GTG CCC ATC TTA TTC		6515	
40	Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe			
	2130	2135	2140	2145
	CTG ATA CTG CTG AGC CTG GGG CTG GGG TTT GCC ATC CTG TAC ACG AAG		6563	
45	Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys			
	2150	2155	2160	
50	CAC CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC		6611	
	His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr			
55				

	2165	2170	2175	
5	AGC TCC AGG CTG GGG TCC GCA ATC TTC TCC TCT GGG GAT GAC CTG GGG			6659
	Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly			
10	2180	2185	2190	
	GAA GAT GAT GAA GAT GCC CCT ATG ATA ACT CGA TTT TCA GAT GAC GTC			6707
	Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val			
15	2195	2200	2205	
	CCC ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT CGTGAAATA			6762
	Pro Met Val Ile Ala			
20	2210			
	TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAAGA TGCACTTTGA GTTGCAATAT			6822
25	GTATTTTTA TATGGGCCAA A			6843

30

35

40

45

50

55

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

- (A) NAME: KOWA CO., LTD.
- (B) STREET: 6-29, Nishiki 3-chome, Naka-ku, Nagoya-shi,
- (C) CITY: Aichi
- (E) COUNTRY: Japan
- (F) POSTAL CODE (ZIP): none

(ii) TITLE OF INVENTION: NOVEL LDL RECEPTOR ANALOG PROTEIN AND THE GENE CODING THEREFOR

(iii) NUMBER OF SEQUENCES: 7

(iv) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6639 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA to mRNA

(i) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGGCCACAC	33AGGCAGCAG	GAGGGAGTCG	CGA3TCCCCT	TCCTATTACAC	CCT3GTGCGC	60
CTGCTGCCGC	CCGGGGCTCT	CTGCGAGGTG	TGGACGCGGA	CACTGCACGG	CGGCCGCGCG	120
CCCTTACCCC	AGGAGCGGGG	CTTCCGCGTG	GTGCAGGGCG	ACCCGCGCGA	GCTGCGGCTG	180
TGGGAGCCCG	GGGATGCCAG	GGGGGCGAGC	CGGGCGGACG	AGAAGCCGCT	CCGGAGGAGA	240
CGGAGCGCTG	CCCTGCAGCC	CGAGCCCATC	AAGGTGTACG	GACAGGTCAAG	CCTCAATGAT	300
TCCCCACAATC	AGATGGTGGT	GCACCTGGCC	GGAGAGAAAA	GCAACGTGAT	CGTGGCCTTG	360
GCCCCGGACA	GCCTGGCGTT	GGCCAGGCC	AGGAGCAGTG	ATGTGTACGT	GTCTTATGAC	420
TATGGAAAAT	CATTCAATAA	GATTTCAGAG	AAATTGAACT	TCGGCGCGGG	AAATAACACA	480
GAGGCTGTGG	TGGCCCAGTT	CTACCAACAGC	CCTGCGGACA	ACAAACGGTA	CATCTTCGCA	540
GATGCCTACG	CCCACTATCT	CTGGATCAGC	TTTGACTTCT	GCAACACCAT	CCATGGCTTT	600
TCCATCCCCT	TCCGGGCAGC	TGATCTCCTA	CTCCACAGTA	AGGCCTCCAA	CCTTCTCCTG	660
GGCTTCGACA	GGTCTCACCC	CAACAAGCAG	CTGTGGAAGT	CGGATGATTT	TGGCCAGACC	720
TGGATCATGA	TTCAAGAACCA	CGTGAAGTCC	TTTTCTTGGG	GAATTGATCC	CTATGACAAA	780

	CCAAACACCA TCTACATCGA ACGGCACGAA CCTT TGG-CT ACTCCACGCT TTTCCGAAGT	840
5	ACAGACTTCT TCCAGTCCCG CGAAAACCAG GAAGTGATCT TGGAGGAAGT GAGAGACTTT	900
	CAGCTTCGGG ACAAGTACAT GTTGCTACA AAGGTGGTGC A1 CTCTTGGG CAGTCCACTG	960
	CAGTCTTCTG TCCAGCTCTG GGTCTCTT GGCGGAAAGC CCATGCGGGC CGCCCAGTTT	1020
10	GTTACAAGAC ATCCTATCAA CGAATATTAC ATCGCGGATG CCTCGGAGGA CCAGGTGTTT	1080
	GTGTGTGTCA GTCACAGCAA CAACCGCACC AACCTCTACA TCTCGGAGGC AGAGGGCTTG	1140
	AAGTTCTCTC TGTCCTTGGG GAACGTGCTC TACTACACCC CGGGAGGGGC CGGCAGTGAC	1200
	ACCTTGGTGA GGTACTTTGC AAATGAACCG TTTGCTGACT TCCATCGTGT GGAAGGGTTG	1260
15	CAGGGAGTCT ACATTGCTAC TCTGATTAAT GGTTCTATGA ATGAGGAGAA CATGAGATCT	1320
	GTCATCACCT TTGACAAAGG GGGCACCTGG GAATTTCTGC AGGCTCCAGC CTTCACGGGG	1380
	TATGGAGAGA AAATCAACTG TGAGCTGTCC GAGGGCTGTT CCCTCCACCT GGGCCAGCGC	1440
20	CTCAGCCAGC TGCTCAACCT CCAGCTCCGG AGGATGCCA TCCTGTCAA GGAGTCGGCG	1500
	CCTGGCCTCA TCATTGCCAC GGGCTCAGTG GGAAAGAACT TGGCTAGCAA GACAAACGTG	1560
	TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCTGGACC TCACTACTAT	1620
25	ACATGGGAG ACCATGGCGG CATCATCATG GCCATTGCC AAGGCATGGA AACCAACGAA	1680
	CTGAAGTACA GTACCAACGA AGGGGAGACC TGGAAAGCCT TCACCTTCTC TGAGAAGCCC	1740
	GTGTTTGTGT ATGGGCTCCT CACGGAAACCC GGCAGAGAAGA GCACGGTCTT CACCATCTTT	1800
30	GGCTCCAACA AGGAGAACGT GCACAGCTGG CTCATCCTCC AGGTCAATGC CACAGACGCC	1860
	CTGGGGGTTTC CTTGCACAGA GAACGACTAC AAGCTCTGGT CACCATCTGA TGAGCGGGGG	1920
	AATGAGTGT TGCTTGACA CAAGACTGTT TTCAAACGGA GGACCCCGCA CGCCACATGC	1980
35	TTTAACGGAG AAGACTTTGA CAGGCCGGTG GTTGTGTCCA ACTGCTCCTG CACCCGGAG	2040
	GACTATGAGT GTGACTTTGG CTTCCGGATG AGTGAAGACT TGGCATTAGA GGTGTGTGTT	2100
	CCAGATCCAG GATTTCTGG AAAGTCTCC CCTCCAGTGC CTTGTCCCGT GGGCTCTACG	2160
40	TACAGGCAT CAAGAGGCTA CCGGAAGATT TCTGGGACA CCTGTAGTGG AGGAGATGTT	2220
	GAGGCACCCG TAGAAGGAGA GCTGGTCCCC TGCCCCCTGG CACAAGAGAA CGAGTTCATC	2280
	CTGTACGCCA CGCGCAAGTC CATCCACCGC TATGACCTGG CTTCCGGAAC CACGGAGCAG	2340
	TTGCCCTCA CTGGGTTGCG GGCAGCAGTG GCCCTGGACT TTGACTATGA GCACAACTGC	2400
45	CTGTATTGGT CTGACCTGGC CTTGGACGTC ATCCAGCGCC TCTGTTTGAA CGGGAGTACA	2460
	GGACAAGAGG TGATCATCAA CTCTGACCTG GAGACGGTAG AAGCTTTGGC TTTTGAACCC	2520
	CTCAGCCAAT TACTTACTG GGTGGACGCA GGCTTTAAAA AGATCGAGGT AGCCAATCCA	2580
50	GATGGTGAAT TCCGACTCAC CGTCGTCAAT TCCTGGTGC TGGATCGGCC CGGGCCCTG	2640
	GTCCTTGTC CCCAAGAAGG GATCATGTTC TGGACCGACT GGGGAGACCT GAAGCCTGGG	2700

5	ATTATATCGGA GCAACATGGAA CGGATCTGCC GCCTATCCCC TCGTGTGGAA GGATGTGAAG	1760
	TGGCCCAATG GCATTTCCGT GGACGATCAG TGGATCTACT GGCGGATGC CTACCTGGAC	2820
	TGCATTGAGC GCATCACGTT CAGCGGCCAG CAGCGCTCCG TCATCCTGGAA CAGACTCCCG	2880
	CACCCCTATG CCATTGCTGT CTTTAAGAAT GAGATTTACT GGGATGACTG GTCACAGCTC	2940
10	AGCATATTCC GAGCTTCTAA GTACAGCGGG TCCCAGATGG AGATTCTGGC CAGCCAGCTC	3000
	ACGGGGCTGA TGGACATGAA GATCTTCTAC AAGGGGAAGA ACACAGGAAG CAATGCGTGT	3060
	GTACCCAGGC CGTGCAGCCT GCTGTGCCTG CCCAGAGCCA ACAACAGCAA AAGCTGCAGG	3120
	TGTCAGATG CGCTGGCCAG CAGTGTCCCT CCTTCCGGGG ACCTGATGTG TGACTGCCCT	3180
15	AAGGGCTACG AGCTGAAGAA CAACACGTGT GTCAAAGAAG AAGACACCTG TCTGCGAAC	3240
	CAGTACCGCT GCAGCAACGG CAACTGCATC AACAGCATCT GGTGGTGCAGA TTTGACAAAC	3300
	GAETGGGGAG ACATGAGCGA CGAGAAGAAC TGCCCTACCA CCATCTGCAG CCTGGACACC	3360
20	CAGTCCCGTT GCCAGGAGTC TGGGACGTGC ATCCCGCTCT CCTACAAATG TGACCTCGAG	3420
	GATGACTGTG GGGACAACAG TGACGAAAGG CACTGTGAAA TGACACCAGTG CGGGAGCGAC	3480
	GAATACAACT GCAGCTCGGG CATGTGCATC CGCTCCCTCT GGGTGTGCAG CGGGGACAAAC	3540
25	GAETGCAGGG ACTGGTCCGA CGAGGCCAAC TGCACAGCCA TCTATCACAC CTGTGAGGCC	3600
	TCCAACCTCC AGTGCCGCAA CGGGCACTGC ATCCCCCAGC GGTGGGCGTG TGACGGCGAC	3660
	GCCGACTGCC AGGATGGCTC TGATGAGGAT CCAGCCAAT GTGAGAAGAA GTGCAACGGC	3720
30	TTCCGCTGCC CGAACGGCAC CTGCATTCCC TCCACCAAGC ACTGTGACGG CCTGCACGAT	3780
	TGCTCGGACG GCTCCGACGA GCAGCACTGC GAGCCCCCTGT GTACACGGTT CATGGACTTC	3840
	GTCTGTAAGA ACCGCCAGCA GTGCCCTCTC CACTCCATGG TGTGCGATGG GATCATCCAG	3900
35	TGCCGTGACG GCTCCGACGA GGACCCAGCC TTTGCAGGAT GCTCCCGAGA CCCCAGTTTC	3960
	CACAAGGTGT CGATGAGTT CGGCTTCCAG TGTCAGAACG GCGTGTGCAT CAGCTTGATC	4020
	TGGAAGTGCAG CGGGGATGGA TGACTGCCGG GACTACTCCG ACGAGGCCAA CTGTGAAAAC	4080
40	CCCCACAGAAG CCCCCAACTG CTCCCGCTAC TTCCAGTTCC GGTGTGACAA TGGCCACTGC	4140
	ATCCCCAACAA GGTGGAAGTG TGACAGGGAG AATGACTGTG GGGACTGGTC CGACGAGAAG	4200
	GACTGTGGAG ATTACACATGT ACTTCCGTCT ACGACTCCTG CACCCCTCCAC GTGTCTGCC	4260
45	AATTACTACC GCTGCGGCCGG GGGGGCCTGC GTGATAGACA CGTGGGTTTG TGACGGGTAC	4320
	CGAGATTGCG CAGATGGATC CGACGAGGAA GCCTGCCCT CGCTCCCCAA TGTCACTGCC	4380
	ACCTCCCTCC CCTCCCAAGCC TGGACGATGC GACCGATTG AGTTTGAGTG CCACCAAGCCA	4440
	AAGAAGTGCAG TCCCTAACTG GAGACGCTGT GACGGCCATC AGGATTGCCA GGATGGCCAG	4500
50	GACGAGGCCA ACTGCCAAC TCACAGCACC TTGACCTGCA TGAGCTGGGA GTTCAAGTGT	4560
	GAGGATGGCG AGGCCTGCAT CGTGTGTCA GAACGCTGCG ACGGCTTCCCT GGACTGCTCA	4620

	GATGAGAGCC ACGAGAAGGC CTGCAGTGAT GAGTTAACTG TATACAAAGT ACAGAATCTT	4680
5	CAGTGGACAG CTGACTTCTC TGGGAATGTC ACTTTGACCT GGATGCGGCC CAAAAAAATG	4740
	CCCTCTGCTG CTTGTGTATA CAACGTGTAC TATAGAGTTG TTGGAGAGAG C TATGGAAG	4800
	ACTCTGGAGA CTCACAGCAA TAAGACAAAC ACTGTATTAA AAGTGTGAA ACCAGATACC	4860
10	ACCTACCAGG TTAAAGTGCA CGTTCAGTGC CTGAGCAAGG TGACACAAACAC CAATGACTTT	4920
	GTGACCTTGA GAACTCCAGA GGGATTGCCA GACGCCCTC AGAACCTCCA GCTGTCGCTC	4980
	CACGGGGAAAG AGGAAGGTGT GATTGTGGGC CACTGGAGCC CTCCCACCCA CACCCACGGC	5040
15	CTCATTGCG AATACATTGT AGAGTATAGC AGGAGTGGTT CCAAGGTGTG GACTTCAGAA	5100
	AGGGCTGCTA GTAACTTTAC AGAAATAAAAG AACTTGTGTC TCAACACCCCT GTACACCGTC	5160
	AGAGTGGCTG CGGTGACGAG TCGTGGGATA GGAAACTGGA GCGATTCCAA ATCCATTACC	5220
	ACCGTGAAG GAAAAGCGAT CCCGCCACCA AATATCCACA TTGACAACTA CGATGAAAAT	5280
20	TCCCTGAGTT TTACCTGAC CGTGGATGGG AACATCAAGG TGAATGGCTA TGTGGTGAAC	5340
	CTTTTCTGGG CATTGACAC CCACAAACAA GAGAAGAAAA CCATGAACCTT CCAAGGGAGC	5400
	TCAGTGTCCC ACAAAAGTTGG CAATCTGACA GCACAGACGG CCTATGAGAT TTCCGCCTGG	5460
25	GCCAAGACTG ACTTGGCGA TAGTCCTCTG TCATTTGAGC ATGTCACGAC CAGAGGGTT	5520
	CGCCACCTG CTCCCTAGCCT CAAGGCCAGG GCTATCAATC AGACTGCACT GGAATGCCACC	5580
	TGGACAGGCC CCAGGAATGT GGTGTATGGC ATTTTCTATG CCACATCCTT CCTGGACCTC	5640
30	TACCGCAACC CAAGCAGCCT GACCACGCCG CTGCACAAACG CAACCGTGCT CGTCGGTAAG	5700
	GATGAGCAGT ATCTGTTCT GGTCCGGGTG GTGATGCCCT ACCAAGGGCC GTCCCTGGAC	5760
	TACGTGGTCG TGAAGATGAT CCCGGACAGC AGGCTTCCTC CCCGGCACCT GCATGCCGTT	5820
35	CACACCGGCA AGACCTCGGC CGTCATCAAG TGGGAGTCGC CCTACGACTC TCCTGACCAG	5880
	GACCTGTTCT ATGCGATCGC AGTTAAAGAT CTGATACGAA AGACGGACCG GAGCTACAAA	5940
	GTCAAGTCCC GCAACAGCAC CGTGGAGTAC ACCCTGAGCA AGCTGGAGCC CGGAGGGAAA	6000
40	TACCACTGCA TTGTGCAGCT GGGGAACATG AGCAAAGATG CCAGTGTGAA GATCACCACC	6060
	GTTCGTTAT CGGCACCCGA TGCCTTAAAA ATCATAACAG AAAATGACCA CGTCCTTCTC	6120
	TTCTGGAAAA GTCTAGCTCT AAAGGAAAAG TATTTAACG AAAGCAGGG CTACGAGATA	6180
45	CACATGTTTG ATAGCGCCAT GAATATCACC GCATACCTTG GGAATACTAC TGACAATTTC	6240
	TTTAAAATTT CCAACCTGAA GATGGGTAC AATTACACAT TCACGGTCCA GGCACGATGC	6300
	CTTTTGGGCA GCCAGATCTG CGGGGAGCCT GCCGTGCTAC TGTATGATGA GCTGGGTCT	6360
50	GGTGGCGATG CGTCGGCGAT GCAGGCTGCC AGGTCTACTG ATGTCGCCGC CGTGGTGGTG	6420
	CCCACCTGT TTCTGATACT GCTGAGCCTG GGGGTCGGGT TTGCCATCCT GTACACGAAG	6480
	CATCGGAGGC TGCAGAGCAG CTTCACCGCC TTGCGCAACA GCCACTACAG CTCCAGACTC	6540

GGCTCCGCCA TCTTCTCCTC TGGGGATGAC TTGGGGAGG ATGATGAAGA TGCTCCTATG 600
 ATCACTGGAT TTTCGGACGA CGTCCCCATG GTGATAGCC 6639

5 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2213 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS:
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
 1 5 10 15

20 Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
 20 25 30

Arg Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe
 35 40 45

25 Arg Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly
 50 55 60

Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg
 65 70 75 80

30 Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
 85 90 95

Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
 100 105 110

35 Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
 115 120 125

Arg Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
 130 135 140

40 Phe Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr
 145 150 155 160

Glu Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
 165 170 175

45 Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
 180 185 190

Phe Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
 195 200 205

50 Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
 210 215 220

Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
 225 230 235 240

55

Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
 245 250 255
 Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
 260 265 270
 Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
 275 280 285
 Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp
 290 295 300
 Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu
 305 310 315 320
 Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
 325 330 335
 Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
 340 345 350
 Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
 355 360 365
 Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
 370 375 380
 Ser Leu Glu Asn Val Leu Tyr Tyr Pro Gly Gly Ala Gly Ser Asp
 385 390 395 400
 Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
 405 410 415
 Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
 420 425 430
 Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
 435 440 445
 Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
 450 455 460
 Ile Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg
 465 470 475 480
 Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
 485 490 495
 Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
 500 505 510
 Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
 515 520 525
 Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp
 530 535 540
 His Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
 545 550 555 560
 Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe
 565 570 575

Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
 580 585 590
 5 Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
 595 600 605
 Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
 610 615 620
 10 Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly
 625 630 635 640
 Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro
 645 650 655
 His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val
 660 665 670
 15 Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe
 675 680 685
 Arg Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly
 20 690 695 700
 Phe Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr
 705 710 715 720
 25 Tyr Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser
 725 730 735
 Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro
 740 745 750
 Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile
 30 755 760 765
 His Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr
 770 775 780
 35 Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys
 785 790 795 800
 Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu
 805 810 815
 Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr
 40 820 825 830
 Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val
 835 840 845
 Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe
 45 850 855 860
 Arg Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu
 865 870 875 880
 Val Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp
 50 885 890 895
 Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr
 900 905 910
 Arg Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp

55

	915	920	925
	Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg		
5	930 935 940		
	Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro		
	945 950 955 960		
10	His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp		
	965 970 975		
	Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln		
	980 985 990		
15	Met Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile		
	995 1000 1005		
	Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro		
	1010 1015 1020		
	Cys Ser Leu Leu Cys Leu Pro Arg Ala Asn Asn Ser Lys Ser Cys Arg		
	1025 1030 1035 1040		
20	Cys Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met		
	1045 1050 1055		
	Cys Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys		
	1060 1065 1070		
25	Glu Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn		
	1075 1080 1085		
	Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp		
	1090 1095 1100		
30	Met Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr		
	1105 1110 1115 1120		
	Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys		
	1125 1130 1135		
35	Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys		
	1140 1145 1150		
	Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met		
	1155 1160 1165		
40	Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp		
	1170 1175 1180		
	Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala		
	1185 1190 1195 1200		
45	Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala		
	1205 1210 1215		
	Cys Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala		
	1220 1225 1230		
	Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys		
	1235 1240 1245		
50	Ile Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly		
	1250 1255 1260		

Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe
 1265 1270 1275 1280
 5 Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp
 1285 1290 1295
 Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala
 1300 1305 1310
 10 Gly Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly
 1315 1320 1325
 Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp
 1330 1335 1340
 15 Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn
 1345 1350 1355 1360
 Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp
 1365 1370 1375
 20 Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp
 1380 1385 1390
 Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu
 1395 1400 1405
 Pro Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg
 1410 1415 1420
 25 Cys Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr
 1425 1430 1435 1440
 Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro
 1445 1450 1455
 30 Asn Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg
 1460 1465 1470
 Phe Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg
 1475 1480 1485
 35 Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn
 1490 1495 1500
 Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys
 1505 1510 1515 1520
 40 Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe
 1525 1530 1535
 Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu
 1540 1545 1550
 45 Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly
 1555 1560 1565
 Asn Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala
 1570 1575 1580
 50 Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys
 1585 1590 1595 1600

55

Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu
 1605 1610 1615
 Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser
 1620 1625 1630
 Lys Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly
 1635 1640 1645
 Leu Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu
 1650 1655 1660
 Glu Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly
 1665 1670 1675 1680
 Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val
 1685 1690 1695
 Trp Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu
 1700 1705 1710
 Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg
 1715 1720 1725
 Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly
 1730 1735 1740
 Lys Ala Ile Pro Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn
 1745 1750 1755 1760
 Ser Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly
 1765 1770 1775
 Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys
 1780 1785 1790
 Leu Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn
 1795 1800 1805
 Leu Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp
 1810 1815 1820
 Leu Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val
 1825 1830 1835 1840
 Arg Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala
 1845 1850 1855
 Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe
 1860 1865 1870
 Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr
 1875 1880 1885
 Thr Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr
 1890 1895 1900
 Leu Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp
 1905 1910 1915 1920
 Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His
 1925 1930 1935
 Leu His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu

55

	1940	1945	1950
5	Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val 1955 1960 1965		
	Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg 1970 1975 1980		
10	Asn Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys 1985 1990 1995 2000		
	Tyr His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val 2005 2010 2015		
15	Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile 2020 2025 2030		
	Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys 2035 2040 2045		
20	Glu Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp 2050 2055 2060		
	Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val 2085 2090 2095		
25	Gln Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val 2100 2105 2110		
	Leu Leu Tyr Asp Glu Leu Gly Ser Gly Gly Asp Ala Ser Ala Met Gln 2115 2120 2125		
30	Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe 2130 2135 2140		
	Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys 2145 2150 2155 2160		
35	His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr 2165 2170 2175		
	Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly 2180 2185 2190		
40	Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val 2195 2200 2205		
	Pro Met Val Ile Ala 2210		

(2) INFORMATION FOR SEQ ID NO: 3:

45 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6961 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50 (ii) MOLECULE TYPE: cDNA to mRNA

55

(ix) FEATURE:
 (A) NAME/KEY: sig peptide
 (B) LOCATION: 178..261

(ix) FEATURE:
 (A) NAME/KEY: mat peptide
 (B) LOCATION: 262..6816

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

10	CCGCGAGCCG CACACGTGAC GGCGCCGCGC CGCGCCGCGC CGCGCCGAGC GGGACCCAGC	60
	GGCTGCCCGG AGCCCCGGGA CGGGCGCGCG CGCGGCCCGC GCCCCGCCGC TCGGCCGGCG	120
	GCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GCGCGCTTCG CCCGAACATG	180
	Met	1
15	GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCC TTC CTA TTC ACC	228
	Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr	
	5 10 15	
	CTG GTC GCG CTG CTG CCG CCC GGG GCT CTC TGC GAG GTG TGG ACG CGG	276
	Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Arg	
	20 25 30	
20	ACA CTG CAC GGC GGC CGC GCG CCC TTA CCC CAG GAG CGG GGC TTC CGC	324
	Thr Leu His Gly Gly Arg Ala Pro Leu Pro Gln Glu Arg Gly Phe Arg	
	35 40 45	
	GTG GTG CAG GGC GAC CCG CGC GAG CTG CGG CTG TGG GAG CGC GGG GAT	372
	Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Glu Arg Gly Asp	
	50 55 60	
25	GCC AGG GGG GCG AGC CGG GCG GAC GAG AAG CCG CTC CGG AGG AGA CGG	420
	Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Arg Arg	
	70 75 80	
	AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTC AGC	468
	Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser	
	85 90 95	
30	CTC AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCC GGA GAG AAA	516
	Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys	
	100 105 110	
	AGC AAC GTG ATC GTG GCC TTG GCC CGG GAC AGC CTG GCG TTG GCC AGG	564
	Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg	
	115 120 125	
35	CCC AGG AGC AGT GAT GTG TAC GTG TCT TAT GAC TAT GGA AAA TCA TTC	612
	Pro Arg Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe	
	130 135 140 145	
	AAT AAG ATT TCA GAG AAA TTG AAC TTC GGC GCG GGA AAT AAC ACA GAG	660
	Asn Lys Ile Ser Glu Lys Leu Asn Phe Gly Ala Gly Asn Asn Thr Glu	
	150 155 160	
40	GCT GTG GTG GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAA CGG TAC	708
	Ala Val Val Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr	
	165 170 175	
	ATC TTC GCA GAT GCC TAC GCC CAG TAT CTC TGG ATC ACG TTT GAC TTC	756
	Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe	
	180 185 190	
45	TGC AAC ACC ATC CAT GGC TTT TCC ATC CCG TTC CGG GCA GCT GAT CTC	804
	Cys Asn Thr Ile His Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu	
	195 200 205	
	CTA CTC CAC AGT AAG GCC TCC AAC CTT CTC CTG GGC TTC GAC AGG TCT	852
	Leu Leu His Ser Lys Ala Ser Asn Leu Leu Gly Phe Asp Arg Ser	
	210 215 220 225	
50	CAC CCC AAC AAG CAG CTG TGG AAG TCG GAT GAT TTT GGC CAG ACC TGG	900
	His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Trp	
	230 235 240	
	ATC ATG ATT CAA GAA CAC GTG AAG TCC TTT TCT TGG GGA ATT GAT CCC	948
	Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro	
	245 250 255	

	TAT GAC AAA CCA AAC ACC ATC TAC ATC GAA CGG CAC GAA CCT TCT GGC	996
	Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly	
	260 265 270	
5	TAC TCC ACG GTT TTC CGA AGT ACA GAC TTC TTC CAG TCC CGG GAA AAC	1044
	Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn	
	275 280 285	
	CAG GAA GTG ATC TTG GAG GAA GTG AGA GAC TTT CAG CTT CGG GAC AAG	1092
	Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys	
	290 295 300 305	
10	TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC AGT CCA CTG CAG	1140
	Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Pro Leu Gln	
	310 315 320	
	TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG CGG GCC	1188
	Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala	
	325 330 335	
15	GCC CAG TTT GTT ACA AGA CAT CCT ATC AAC GAA TAT TAC ATC GCG GAT	1236
	Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp	
	340 345 350	
	GCC TCG GAG GAC CAG GTG TTT GTG TGT GTC AGT CAC AGC AAC AAC CGC	1284
	Ala Ser Gln Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg	
	355 360 365	
20	ACC AAC CTC TAC ATC TCG GAG GCA GAG GGC TTG AAG TTC TCT CTG TCC	1332
	Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser	
	370 375 380 385	
	CTG GAG AAC GTG CTC TAC TAC ACC CCG GGA GGG GCC GGC AGT GAC ACC	1380
	Leu Glu Asn Val Leu Tyr Tyr Thr Pro Gly Gly Ala Gly Ser Asp Thr	
	390 395 400	
25	TTG CTG AGG TAC TTT GCA AAT GAA CCG TTT GCT GAC TTC CAT CGT GTG	1428
	Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val	
	405 410 415	
	CAA CGG TTG CAG GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG	1476
	Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met	
	420 425 430	
30	AAT GAG CAG AAC ATG AGA TCT GTC ATC ACC TTT GAC AAA GGG GGC ACC	1524
	Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr	
	435 440 445	
	TGG GAA TTT CTG CAG GCT CCA GCC TTC ACG GGG TAT GGA GAG AAA ATC	1572
	Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile	
	450 455 460 465	
35	AAC TGT CAG CTG TCC GAG GGC TGT TCC CTC CAC CTG GCC CAG CGC CTC	1620
	Asn Cys Glu Leu Ser Glu Gly Cys Ser Leu His Leu Ala Gln Arg Leu	
	470 475 480	
	AGC CAG CTG CTC AAC CTC CAG CTC CGG AGG ATG CCC ATC CTG TCC AAG	1663
	Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys	
	485 490 495	
40	GAG TCG GCG CCT GGC CTC ATC ATT GCC ACG GGC TCA GTG GGA AAG AAC	1716
	Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys Asn	
	500 505 510	
	TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG	1764
	Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala Arg	
	515 520 525	
45	TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAT ACA TGG GGA GAC CAT	1812
	Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His	
	530 535 540 545	
	GGC GGC ATC ATC ATG GCC ATT GCC CAA GGC ATG GAA ACC AAC GAA CTG	1860
	Gly Gly Ile Ile Met Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu	
	550 555 560	
50	AAG TAC AGT ACC AAC GAA GGG GAG ACC TGG AAA GCC TTC ACC TTC TCT	1903
	Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Ala Phe Thr Phe Ser	
	565 570 575	
	GAG AAG CCC GTG TTT GTG TAT GGG CTC CTC ACG GAA CCC GGC GAG AAG	1956
	Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys	
	580 585 590	
	AGC ACG GTC TTC ACC ATC TTT GGC TCC AAC AAG GAG AAC GTG CAC AGC	2004

5	Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser 595 600 605	2052
	TGG CTC ATC CTC CAG GTC AAT GCC ACA GAC GCC CTG GGG GTT CCT TGC Tip Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys 610 615 620 625	
	ACA GAG AAC GAC TAC AAG CTC TGG TCA CCA TCT GAT GAG CGG GGG AAT Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn 630 635 640 645	2100
10	GAG TGT TTG CTT GGA CAC AAG ACT GTT TTC AAA CGG AGG ACC CCG CAC Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His 645 650 655 660	2148
	GCC ACA TGC TTT AAC GGA GAA GAC TTT GAC AGG CCG GTG GTT GTG TCC Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser 660 665 670 675	2196
15	AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTT GGC TTC CGG Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Arg 675 680 685 690	2244
	ATG AGT GAA GAC TTG GCA TTA GAG GTG TGT GTT CCA GAT CCA GGA TTT Met Ser Glu Asp Leu Ala Leu Glu Val Cys Val Pro Asp Pro Gly Phe 695 700 705 710	2292
20	TCT GGA AAG TCC TCC CCT CCA GTG CCT TGT CCC GTG GGC TCT ACG TAC Ser Gly Lys Ser Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr Tyr 710 715 720 725	2340
	AGG CGA TCA AGA GGC TAC CGG AAG ATT TCT GGG GAC ACC TGT AGT GGA Arg Arg Ser Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser Gly 730 735 740 745	2388
25	GGA GAT GTT GAG GCA CGG CTA GAA GGA GAG CTG GTC CCC TGT CCC CTG Gly Asp Val Glu Ala Arg Leu Glu Gly Leu Val Pro Cys Pro Leu 740 745 750 755	2436
	GCA GAA GAG AAC GAG TTC ATC CTG TAC GCC ACG CGC AAG TCC ATC CAC Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Thr Arg Lys Ser Ile His 760 765 770 775	2484
30	CGC TAT GAC CTG GCT TCC GGA ACC ACG GAG CAG TTG CCC CTC ACT GGG Arg Tyr Asp Leu Ala Ser Gly Thr Thr Glu Gln Leu Pro Leu Thr Gly 780 785 790 795	2532
	TTG CCG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGC CTG Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys Leu 800 805 810 815	2580
35	TAT TGG TCT GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG AAC Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu Asn 820 825 830 835	2628
	GGG AGT ACA GGA CAA GAG GTG ATC ATC AAC TCT GAC CTG GAG ACG GTA Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Asp Leu Glu Thr Val 840 845 850 855	2676
40	GAA GCT TTG GCT TTT GAA CCC CTC AGC CAA TTA CTT TAC TGG GTG GAC Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val Asp 860 865 870 875	2724
	GCA GGC TTT AAA AAG ATC GAG GTA GCC AAT CCA GAT GGT GAC TTC CGA Ala Gly Phe Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe Arg 880 885 890 895	2772
45	CTC ACC GTC GTC AAT TCC TCG GTG CTG GAT CGG CCC CGG GCC CTG GTC Leu Thr Val Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu Val 900 905 910 915	2820
	CTT GTG CCC CAA GAA GGG ATC ATG TTC TGG ACC GAC TGG GGA GAC CTG Leu Val Pro Gln Glu Gly Ile Met Phe Trp Thr Asp Trp Gly Asp Leu 920 925 930 935	2868
50	AAG CCT GGG ATT TAT CGG AGC AAC ATG GAC GGA TCT GCC GCC TAT CGC Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr Arg 940 945 950 955	2916
	CTC GTG TCG GAG GAT GTG AAG TGG CCC AAT GGC ATT TCC GTG GAC GAT Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp Asp 960 965 970 975	2964
	CAG TGG ATC TAC TGG ACG GAT GCC TAC CTG GAC TGC ATT GAG CGC ATC Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Asp Cys Ile Glu Arg Ile	3012

	930	935	940	945	
	ACG TTC AGC GGC CAG CAG CGC TCC GTC ATC TGT GAC ACA CTC CCG CAC				3060
	Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Arg Leu Pro His				
5	950	955	960		
	CCC TAT GCC ATT GCT GTC TTT AAG AAT GAG ATT TAC TGG GAT GAC TGG				3108
	Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp Trp				
	965	970	975		
	TCA CAG CTC AGC ATA TTC CGA GCT TCT AAG TAC AGC GGG TCC CAG ATG				3156
	Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln Met				
10	980	985	990		
	GAG ATT CTG GCC AGC CAG CTC ACG GGG CTG ATG GAC ATG AAG ATC TTC				3204
	Glu Ile Leu Ala Ser Gln Leu Thr Gly Leu Met Asp Met Lys Ile Phe				
	995	1000	1005		
	TAC AAG GGG AAG AAC ACA GGA AGC AAT GCG TGT GTA CCC AGG CCG TGC				3252
	Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro Cys				
15	1010	1015	1020	1025	
	AGC CTG CTG TGC CTG CCC AGA GCC AAC AGC AAA AGC TGC AGG TGT				3300
	Ser Leu Leu Cys Leu Pro Arg Ala Asn Ser Lys Ser Cys Arg Cys				
	1030	1035	1040		
	CCA GAT GGC GTG GCC AGC AGT GTC CTC CCT TCC GGG GAC CTG ATG TGT				3348
	Pro Asp Gly Val Ala Ser Ser Val Leu Pro Ser Gly Asp Leu Met Cys				
	1045	1050	1055		
20	GAC TGC CCT AAG GGC TAC GAG CTG AAG AAC ACG TGT GTC AAA GAA				3396
	Asp Cys Pro Lys Gly Tyr Glu Leu Lys Asn Asn Thr Cys Val Lys Glu				
	1060	1065	1070		
	GAA GAC ACC TGT CTG CGC AAC CAG TAC CGC TGC AGC AAC GGG AAC TGC				3444
	Glu Asp Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn Cys				
	1075	1080	1085		
25	ATC AAC AGC ATC TGG TGG TGC GAT TTC GAC AAC GAC TGC GGA GAC ATG				3492
	Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp Met				
	1090	1095	1100	1105	
	AGC GAC GAG AAG AAC TGC CCT ACC ACC ATC TGC GAC CTG GAC ACC CAG				3540
	Ser Asp Glu Lys Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr Gln				
	1110	1115	1120		
30	TTC CGT TGC CAG GAG TCT GGG ACG TGC ATC CCG CTC TCC TAC AAA TGT				3583
	Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys Cys				
	1125	1130	1135		
	GAC CTC GAG GAT GAC TGT GGG GAC AAC AGT GAC GAA AGG CAC TGT GAA				3636
	Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Arg His Cys Glu				
	1140	1145	1150		
35	ATG CAC CAG TGC CGG AGC GAC GAA TAC AAC TGC AGC TCG GGC ATG TGC				3684
	Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met Cys				
	1155	1160	1165		
	ATC CGC TCC TCC TGG GTG TGC GAC CGG GAC AAC GAC TGC AGG GAC TGG				3732
	Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp Trp				
	1170	1175	1180	1185	
40	TCC GAC GAG GCC AAC TGC ACA GCC ATC TAT CAC ACC TGT GAG GCC TCC				3780
	Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala Ser				
	1190	1195	1200		
	AAC TTC CAG TGC CGC AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG TGT				3828
	Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala Cys				
	1205	1210	1215		
45	GAC GGC GAC GCC GAC TGC CAG GAT CGC TCT GAT GAG GAT CCA GCC AAC				3876
	Asp Gly Asp Ala Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Ala Asn				
	1220	1225	1230		
	TGT GAG AAG AAG TGC AAC GGC TTC CGC TGC CCG AAC GGC ACC TGC ATT				3924
	Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys Ile				
	1235	1240	1245		
50	CCC TCC ACC AAG CAC TGT GAC CGC CTG CAC GAT TGC TCG GAC GGC TCC				3972
	Pro Ser Thr Lys His Cys Asp Gly Leu His Asp Cys Ser Asp Gly Ser				
	1250	1255	1260	1265	
	GAC GAG CAG CAC TGC GAG CCC CTG TGT ACA CGG TTC ATG GAC TTC GTG				4020
	Asp Glu Gln His Cys Glu Pro Leu Cys Thr Arg Phe Met Asp Phe Val				
	1270	1275	1280		

	TGT AAC CCC CAG CAG TGC CTC TTC (AC TCC ATG CTG TG:G GAT GGG Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp Gly 1285 1290 1295	4063
5	ATC ATC CAG TGC CGT GAC GGC TCC GAC GAG GAC CCA GCC TTT GCA GGA Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Pro Ala Phe Ala Gly 1300 1305 1310	4116
	TGC TCC CGA GAC CCC GAG TTC CAC AAG GTG TGC GAT GAG TTC GGC TTC Cys Ser Arg Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly Phe 1315 1320 1325	4164
10	CAG TGT CAG AAC GGC GTG TGC ATC AGC TTG ATC TGG AAG TGC GAC GGG Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp Gly 1330 1335 1340 1345	4212
	ATG GAT GAC TGC GGG GAC TAC TCC GAC GAG GCC AAC TGT GAA AAC CCC Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn Pro 1350 1355 1360	4260
15	ACA GAA GCC CCC AAC TGC TCC CGC TAC TTC CAG TTC CGG TGT GAC AAT Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Asp Asn 1365 1370 1375	4308
	GGC CAC TGC ATC CCC AAC AGG TGG AAG TGT GAC AGG GAG AAT GAC TGT Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp Cys 1380 1385 1390	4356
20	GGG GAC TGG TCC GAC GAG AAG GAC TGT GGA GAT TCA CAT GTC CTT CCG Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Val Leu Pro 1395 1400 1405	4404
	TCT ACG ACT CCT GCA CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC TGC Ser Thr Thr Pro Ala Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg Cys 1410 1415 1420 1425	4452
25	GGC GGG GGG GCC TGC GTG ATA GAC ACG TGG GTT TGT GAC GGG TAC CGA Gly Gly Gly Ala Cys Val Ile Asp Thr Trp Val Cys Asp Gly Tyr Arg 1430 1435 1440	4500
	GAT TGC GCA GAT GGA TCC GAC GAG GAA GCC TGC CCC TCG CTC CCC AAT Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Ser Leu Pro Asn 1445 1450 1455	4548
30	GTC ACT GCC ACC TCC TCC CCC TCC CAG CCT GGA CGA TGC GAC CGA TTT Val Thr Ala Thr Ser Ser Pro Ser Gln Pro Gly Arg Cys Asp Arg Phe 1460 1465 1470	4596
	GAG TTT GAG TGC CAC CAG CCA AAG AAG TGC ATC CCT AAC TGG AGA CGC Glu Phe Glu Cys His Gln Pro Lys Lys Cys Ile Pro Asn Trp Arg Arg 1775 1480 1485	4644
35	TGT GAC GGC CAT CAG GAT TGC CAG GAT GGC CAG GAC GAG GCC AAC TGC Cys Asp Gly His Gln Asp Cys Gln Asp Gly Gln Asp Glu Ala Asn Cys 1490 1495 1500 1505	4692
	CCC ACT CAC AGC ACC TTG ACC TGC ATG AGC TGG GAG TTC AAG TGT GAG Pro Thr His Ser Thr Leu Thr Cys Met Ser Trp Glu Phe Lys Cys Glu 1510 1515 1520	4740
	GAT GGC GAG GCC TGC ATC GTG CTG TCA GAA CGC TGC GAC GGC TTC CTG Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe Leu 1525 1530 1535	4788
40	GAC TGC TCA GAT GAG AGC GAC GAG AAG GCC TGC AGT GAT GAG TTA ACT Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu Thr 1540 1545 1550	4836
	GTA TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG AAT Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly Asn 1555 1560 1565	4884
45	GTC ACT TTG ACC TGG ATG CGG CCC AAA AAA ATG CCC TCT GCT GCT TGT Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ala Cys 1570 1575 1580 1585	4932
	GTA TAC AAC GTG TAC TAT AGA GTT GGT GGA GAG AGC ATA TGG AAG ACT Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys Thr	4980
50	1590 1595 1600	5028
	CTG GAG ACT CAC AGC AAT AAG ACA AAC ACT GTC TTA AAA GTG TTG AAA Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu Lys 1605 1610 1615	5076
	CCA GAT ACC ACC TAC CAG GTT AAA GTG CAG GTT CAG TGC CTG AGC AAG	

5	Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Lys Ile Ser Lys 1620 1625 1630	5124
	GTG CAC AAC ACC AAT GAC TTT GTG ACC TTG AGA ACT CCA GAG GGA TTG Val His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly Leu 1635 1640 1645	
	CCA GAC GCC CCT CAG AAC CTC CAG CTG TCG CTC CAC GGG GAA GAG GAA Pro Asp Ala Pro Gln Asn Leu Gln Leu Ser Leu His Gly Glu Glu Glu 1650 1655 1660 1665	5172
	GGT GTG ATT GTG GGC CAC TGG AGC CCT CCC ACC CAC ACC CAC GGC CTC Gly Val Ile Val Gly His Trp Ser Pro Pro Thr His Thr His Gly Leu 1670 1675 1680	5220
	ATT CGC GAA TAC ATT GTA GAG TAT AGC AGG AGT GGT TCC AAG GTG TGG Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Val Trp 1685 1690 1695	5268
10	ACT TCA GAA AGG GCT GCT AGT AAC TTT ACA GAA ATA AAG AAC TTG TTG Thr Ser Glu Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu Leu 1700 1705 1710	5316
	GTC AAC ACC CTG TAC ACC GTC AGA GTG GCT GCG GTG ACG AGT CGT GGG Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg Gly 1715 1720 1725	5364
	ATA GGA AAC TGG AGC GAT TCC AAA TCC ATT ACC ACC GTG AAA GGA AAA Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Val Lys Gly Lys 1730 1735 1740 1745	5412
20	GCG ATC CCG CCA CCA AAT ATC CAC ATT GAC AAC TAC GAT GAA AAT TCC Ala Ile Pro Pro Asn Ile His Ile Asp Asn Tyr Asp Glu Asn Ser 1750 1755 1760	5460
	CTG AGT TTT ACC CTG ACC GTG GAT GGG AAC ATC AAG GTG AAT GGC TAT Leu Ser Phe Thr Leu Thr Val Asp Gly Asn Ile Lys Val Asn Gly Tyr 1765 1770 1775	5508
	GTG GTG AAC CTT TTC TCG GCA TTT GAC ACC CAC AAA CAA GAG AAG AAA Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Lys Lys 1780 1785 1790	5556
25	ACC ATG AAC TTC CAA GGG AGC TCA GTG TCC CAC AAA GTT GGC AAT CTG Thr Met Asn Phe Gln Gly Ser Ser Val Ser His Lys Val Gly Asn Leu 1795 1800 1805	5604
	ACA GCA CAG ACG GCC TAT GAG ATT TCC GCC TGG GCC AAG ACT GAC TTG Thr Ala Gln Thr Ala Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp Leu 1810 1815 1820 1825	5652
	GGC GAT AGT CCT CTG TCA TTT GAG CAT GTC ACG ACC AGA GGG GTT CGC Gly Asp Ser Pro Leu Ser Phe Glu His Val Thr Thr Arg Gly Val Arg 1830 1835 1840	5700
30	CCA CCT GCT CCT AGC CTC AAG GCC AGG GCT ATC AAT CAG ACT GCA GTG Pro Pro Ala Pro Ser Leu Lys Ala Arg Ala Ile Asn Gln Thr Ala Val 1845 1850 1855	5748
	GAA TGC ACC TGG ACA GGC CCC AGG AAT GTG GTG TAT GGC ATT TTC TAT Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe Tyr 1860 1865 1870	5796
40	GGC ACA TCC TTC CTG GAC CTC TAC CGC AAC CCA AGC AGC CTG ACC ACG Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Ser Ser Leu Thr Thr 1875 1880 1885	5844
	CCG CTG CAC AAC GCA ACC GTG CTC GTC GGT AAG GAT GAG CAG TAT CTG Pro Leu His Asn Ala Thr Val Leu Val Gly Lys Asp Glu Gln Tyr Leu 1890 1895 1900 1905	5892
	TTT CTG GTC CGG GTG GTG ATG CCC TAC CAA GGG CCG TCC TCG GAC TAC Phe Leu Val Arg Val Val Met Pro Tyr Gln Gly Pro Ser Ser Asp Tyr 1910 1915 1920	5940
45	GTG STC GTG AAG ATG ATC CCG GAC AGC AGG CTT CCT CCC CGG CAC CTG Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His Leu 1925 1930 1935	5988
	CAT GCC GTT CAC ACC GGC AAG ACC TCG GCC GTC ATC AAG TGG GAG TCG His Ala Val His Thr Gly Lys Thr Ser Ala Val Ile Lys Trp Glu Ser 1940 1945 1950	6036
50	CCC TAC GAC TCT CCT GAC CAG GAC CTG TTC TAT GCG ATC GCA GTT AAA Pro Tyr Asp Ser Pro Asp Gln Asp Leu Phe Tyr Ala Ile Ala Val Lys	6084

	1955	1960	1965	
	GAT CTG ATA CGA AAG ACG GAC CGG AGC TAC AAA	GTC AAG TCC CGC AAC		6132
5	Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg Asn			
	1970	1975	1980	1985
	AGC ACC GTG GAG TAC ACC CTG AGC AAG CTG GAG CCC GGA GGG AAA TAC			6180
	Ser Thr Val Glu Tyr Thr Leu Ser Lys Leu Glu Pro Gly Gly Lys Tyr			
	1990	1995	2000	
	CAC GTC ATT GTG CAG CTG GGG AAC ATG AGC AAA GAT GCC AGT GTG AAG			6228
10	His Val Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ala Ser Val Lys			
	2005	2010	2015	
	ATC ACC ACC GTT TCG TTA TCG GCA CCC GAT GCC TTA AAA ATC ATA ACA			6276
	Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile Thr			
	2020	2025	2030	
	GAA AAT GAC CAC GTC CTT CTC TTC TGG AAA AGT CTA GCT CTA AAG GAA			6324
15	Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys Glu			
	2035	2040	2045	
	AAG TAT TTT AAC GAA AGC AGG GGC TAC GAG ATA CAC ATG TTT GAT AGC			6372
	Lys Tyr Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp Ser			
	2050	2055	2060	2065
	GCC ATG AAT ATC ACC GCA TAC CTT GGG AAT ACT ACT GAC AAT TTC TTT			6420
	Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe Phe			
	2070	2075	2080	
20	AAA ATT TCC AAC CTG AAG ATG GGT CAC AAT TAC ACA TTC ACG GTC CAG			6468
	Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val Gln			
	2085	2090	2095	
	GCA CGA TGC CTT TTG GGC AGC CAG ATC TGC GGG GAG CCT GCC GTG CTA			6516
	Ala Arg Cys Leu Leu Gly Ser Gln Ile Cys Gly Glu Pro Ala Val Leu			
	2100	2105	2110	
25	CTG TAT GAT GAG CTG GGG TCT GGT GGC GAT GCG TCG GCG ATG CAG GCT			6564
	Leu Tyr Asp Glu Leu Gly Ser Gly Asp Ala Ser Ala Met Gln Ala			
	2115	2120	2125	
	GCC AGG TCT ACT GAT GTC GCC GCC GTG GTG GTG CCC ATC CTG TTT CTG			6612
	Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu Phe Leu			
	2130	2135	2140	2145
30	ATA CTG CTG AGC CTG GGG GTC GGG TTT GCC ATC CTG TAC ACG AAG CAT			6660
	Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys His			
	2150	2155	2160	
	CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC AGC			6708
	Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr Ser			
	2165	2170	2175	
35	TCC AGA CTC GGC TCC GCC ATC TTC TCC TCT GGG GAT GAC TTG GGG GAG			6756
	Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly Glu			
	2180	2185	2190	
	GAT GAT GAA GAT GCT CCT ATG ATC ACT GGA TTT TCG GAC GAC GTC CCC			6804
	Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val Pro			
	2195	2200	2205	
40	ATG GTG ATA GCC TGAAAGAGCT TTCCCTCACTA GAAACCAAAT GGTGTAAATA			6856
	Met Val Ile Ala			
	2210			
	TTTTATTTGA TAAAGATAGT TGATGGTTTA TTTTAAAGA TGCACTTTGA GTTGCAATAT			6916
	GTTATTTTA TATGGGCCAA AAACAAAAGC AAAAAAAA AAAAAA			6961

(2) INFORMATION FOR SEQ ID NO: 4:

45

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 300 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

50

(ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

5	ATATCCACAT TGACAGCTAT GGTGAAAATT ATCTAAGCTT CACCCCTGACC ATGGAGAGTG	60
	ATATCAAGGT GAATGGCTAT GTGGTGAACC TTTTCTGGC ATTTGACACC CACAAGCAAG	120
	AGAGGAGAAC TTTGAACCTTC CGAGGAAGCA TATTGTCACA CAAAGTTGGC AATCTGACAG	180
10	CTCATACATC CTATGAGATT TCTGCCTGGG CCAAGACTGA CTTGGGGGAT AGCCCTCTGG	240
	CATTGAGCA TGTATGACC AGAGGGGTTG GCCCACCTGC ACCTAGCCTC AAGGCCAAAG	300

(2) INFORMATION FOR SEQ ID NO: 5:

15 (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6642 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: cDNA to mRNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

25	ATGGCGACAC GGAGCAGCAG GAGGGAGTCG CGACTCCCCT TCCTATTAC CCTGGTCGCA	60
	CTGCTGCCGC CGGGAGCTCT CTGCGAAGTC TGGACGCAGA GGCTGCACGG CGGCAGCGCG	120
	CCCTTGCCTT AGGACCGGGG CTTCTCGTG GTGCAGGGCG ACCCGCGCGA GCTGCGGCTG	180
30	TGGGGCGCG GGGATGCCAG GGGGGCGAGC CGCGCGGACG AGAAGCCGCT CGGGAGGAAA	240
	CGGAGCGCTG CCCTGCAGCC CGAGCCCATC AAGGTGTACG GACAGGTTAG TCTGAATGAT	300
	TCCCCACAATC AGATGGTGGT GCACGGGGCT GGAGAGAAAA GCAACGTGAT CGTGGCCTTG	360
35	GCCCCAGATA GCCTGGCATT GGCGAGGCCA AAGAGCAGTG ATGTGTACGT GTCTTACGAC	420
	TATGGAAAAT CATTCAAGAA AATTTCAGAC AAGTTAAACT TTGGCTTGGG AAATAGGAGT	480
	GAAGCTGTTA TGCGCCAGTT CTACCCACAGC CCTGCGGACA ACAAGCGGTA CATCTTTCGA	540
40	GACGCTTATG CCCAGTACCT CTGGATCACG TTTGACTTCT GCAACACTCT TCAAGGCTTT	600
	TCCATCCCAT TTGGGGCAGC TGATCTCCTC CTACACAGTA AGGCCTCCAA CCTTCTCTTG	660
	GGCTTGACA GGTCCCACCC CAACAAGCAG CTGTGGAAGT CAGATGACTT TGGCCAGACC	720
45	TGGATCATGA TTCAGGAACA TGTCAAGTCC TTTTCTTGGG GAATTGATCC CTATGACAAA	780
	CCAAATAACCA TCTACATTGA ACGACACGAA CCCTCTGGCT ACTCCACTGT CTTCCGAAGT	840
	ACAGATTTCT TCCAGTCCCG GGAAAACCGAG GAACTGATCC TTGAGGAAAGT GAGAGATTTT	900
50	CAGCTTCGGG ACAAGTACAT GTTGCTACA AAGGTGGTGC ATCTCTTGGG CAGTGAACAG	960
	CAGTCTTCTG TCCAGCTCTG GGTCTCCTTT GGCCGGAAGC CCATGAGAGC AGCCCAGTTT	1020
	GTCACAAGAC ATCCTATTAA TGAATATTAC ATCGCAGATG CCTCCGAGGA CCAGGTGTTT	1080

	GTGTGTGTCA GCCACAGTAA CAACCGCACC AATTATACA TCTCAGAAGC AGAAGGGCTG	1140
5	AAGTTCTCCC TGTCCTTGGA GAACGTGCTC TATTACAGCC CAGGAGGGC CGGCAGTGAC	1200
	ACCTTGGTGA GGTATTTGC AAATGAACCA TITGCTGACT TCCACCGAGT GGAAGGATTG	1260
	CAAGGAGTCT ACATTGCTAC TCTGATTAAT GGTCTATGA ATGAGGAGAA CATGAGATCG	1320
10	GTCATCACCT TTGACAAAGG GGGAACCTGG GAGTTCTTC AGGCTCCAGC CTTCACGGGA	1380
	TATGGAGAGA AAATCAATTG TGAGCTTCC CAGGGCTGTT CCCTTCATCT GGCTCAGCGC	1440
	CTCAGTCAGC TCCTCAACCT CCAGCTCCGG AGAATGCCA TCCTGTCAA GGAGTCGGCT	1500
	CCAGGCCTCA TCATGCCAC TGGCTCAGTG GGAAAGAACT TGGCTAGCAA GACAAACGTG	1560
15	TACATCTCTA GCAGTGCTGG AGCCAGGTGG CGAGAGGCAC TTCTGGACC TCACTACTAC	1620
	ACATGGGGAG ACCACGGCGG AATCATCACG GCCATTGCC AGGGCATGGA AACCAACGAG	1680
	CTAAAATACA GTACCAATGA AGGGGAGACC TGGAAACAT TCATCTTCTC TGAGAAGCCA	1740
20	GTGTTTGTGT ATGGCCTCCT CACAGAACCT GGGGAGAAGA GCACTGTCTT CACCATCTT	1800
	GGCTCGAAC AAGAGAATGT CCACAGCTGG CTGATCCTCC AGGTCAATGC CACGGATGCC	1860
	TTGGGAGTTC CCTGCACAGA GAATGACTAC AAGCTGTGGT CACCATCTGA TGAGCGGGGG	1920
25	AATGAGTGTG TGCTGGACA CAAGACTGTT TTCAAACGGC GGACCCCCCA TGCCACATGC	1980
	TTCAATGGAG AGGACTTTGA CAGGCCGGTG GTCGTGTCCA ACTGCTCCTG CACCCGGGAG	2040
	GACTATGAGT GTGACTTCGG TTTCAAGATG AGTGAAGATT TGTCATTAGA GGTTTGTGTT	2100
30	CCAGATCCGG AATTCTCTGG AAAGTCATAC TCCCCCTCTG TGCCCTGCC TGTGGTTCT	2160
	ACTTACAGGA GAACGAGAGG CTACCGGAAG ATTTCTGGGG ACACCTGTAG CGGAGGAGAT	2220
	GTTGAAGCGC GACTGGAAGG AGAGCTGGTC CCCTGTCCCC TGGCAGAAGA GAACGAGTTC	2280
35	ATTCTGTATG CTGTGAGGAA ATCCATCTAC CGCTATGACC TGGCCTCGG AGCCACCGAG	2340
	CAGTTGCCTC TCACCGGGCT ACAGGCAGCA GTGGCCCTGG ACTTTGACTA TGAGCACAAAC	2400
	TGTTTGTATT GGTCCGACCT GGCCCTGGAC GTCATCCAGC GCCTCTGTT GAATGGAAGC	2460
	ACAGGGCAAG AGGTGATCAT CAATTCTGGC CTGGAGACAG TAGAAGCTTT GGCTTTGAA	2520
40	CCCCTCAGCC AGCTGTTTA CTGGTAGAT GCAGGCTTCA AAAAGATTGA GGTAGCTAAT	2580
	CCAGATGGCG ACTTCCGACT CACAATCGTC AATTCTCTG TGCTTGATCG TCCCAGGGCT	2640
	CTGGTCCTCG TGCCCCAAGA GGGGGTGATG TTCTGGACAG ACTGGGGAGA CCTGAAGCCT	2700
45	GGGATTTATC GGAGCAATAT GGATGGTTCT GCTGCCTATC ACCTGGTGTC TGAGGATGTG	2760
	AAGTGGCCCA ATGGCATCTC TGTGGACGAC CAGTGGATTG ACTGGACGGA TGCTTACCTG	2820
	GAGTGCATAG AGCGGATCAC GTTCAGTGGC CAGCAGCGCT CTGTCATTCT GGACAAACCTC	2880
50	CCGCACCCCT ATGCCATTGC TGTCTTTAAG AATGAAATCT ACTGGGATGA CTGGTCACAG	2940
	CTCAGCATAT TCCGAGCTTC CAAATACAGT GGGTCCCAGA TGGAGATTCT GGCAAAACAG	3000

5	CTCACGGGC TCATGGACAT GAAGATTTTC TACAA GGGA AGAACAC TGG AAGCUAATSCC	3060
	TGTGTGCCA GGCCATGCAG CCTGCTGTGC CTGCCAAGG CCAACACAG TAGA AGCTGC	3120
	AGGTGTCCAG AGGATGTGTC CAGCAGTGTG CTTCCATCAG GGGACCTGAT GTGTGACTGC	3180
	CCTCAGGGCT ATCAGCTCAA GAACAATACC TGTGTCAAAG AAGAGAACAC CTGTCTTCGC	3240
10	AACCAGTATC GCTGCAGCAA CGGGAACGTG ATCAACAGCA TTTGGTGGTG TGACTTTGAC	3300
	AACGACTGTG GAGACATGAG CGATGAGAGA AACTGCCCTA CCACCATCTG TGACCTGGAC	3360
	ACCCAGTTTC GTTGCCAGGA GTCTGGGACT TGTATCCCAC TGTCTATAA ATGTGACCTT	3420
	GAGGATGACT GTGGAGACAA CAGTGATGAA AGTCATTGTG AAATGCACCA GTGCCGGAGT	3480
15	GACGAGTACA ACTGCAGTTC CGGCATGTGC ATCCGCTCCT CCTGGGTATG TGACGGGGAC	3540
	AACGACTGCA GGGACTGGTC TGATGAAGCC AACTGTACCG CCATCTATCA CACCTGTGAG	3600
	GCCTCCAAC TCCAGTGCCG AAACGGGCAC TGCA CCCCC AGCGGTGGGC GTGTGACGGG	3660
20	GATA CGG ACT GCCAGGATGG TTCCGATGAG GATCCAGTCA ACTGTGAGAA GAAGTGCAAT	3720
	GGATTCCGCT GCCCAAACGG CACTTGATC CCATCCAGCA AACATTGTGA TGGTCTGCCT	3780
	GATTGCTCTG ATGGCTCCGA TGAACAGCAC TGCGAG CCCC TCTGTACGCA CTTCATGGAC	3840
25	TTTGTGTGTA AGAACCGCCA GCAGTGCCTG TTCCACTCCA TGGTCTGTGA CGGAATCATC	3900
	CAGTGCCGCG ACGGGTCCGA TGAGGATGCG CGCTTGCAG GATGCT CCC AGATCCTGAG	3960
	TTCCACAAGG TATGTGATGA GTTCGGTTTC CAGTGT CAGA ATGGAGTGTG CATCAGTTG	4020
30	ATTTGGAAGT GCGACGGGAT GGATGATTGC GGC GATT TATT CTGATGAAGC CAACTGCGAA	4080
	AACCCCACAG AAGCCCCAAA CTGCT CCC GC TACTTCCAGT TT CGGT GTGA GAATGGCCAC	4140
	TGCA AT CCCCA ACAGATGAA ATGTGACAGG GAGAACGACT GTGGGGACTG GTCTGATGAG	4200
35	AAGGATTGTG GAGATT CACA TATTCTTCCC TTCTCGACTC CTGGGCCCTC CACGTGTCTG	4260
	CCCAATTACT ACCGCTGCAG CAGTGGGACC TGCGT GAT GG ACACCTGGGT GTGC GAC GGG	4320
	TACCGAGATT GTGCAGATGG CTCTGACGAG GAAGCCTGCC CCTTGCT TG C AAACGTCACT	4380
	GCTGCCTCCA CTCCCACCA ACTTGGCGA TGTGACCGAT TTGAGTTCGA ATGCCACCAA	4440
40	CCGAAGACGT GTATT CCC AA CTGGAAGCGC TGTGACGGCC ACCAAGATTG CCAGGATGGC	4500
	CGGGACGAGG CCAATTGCC C CACACACAGC ACCTTGACTT GCATGAGCAG GGAGTTCCAG	4560
	TGGCAGGACG GGGAGGCCTG CATTGTGCTC TCGGAGCGCT GCGACGGCTT CCTGGACTGC	4620
45	TCGGACGAGA GCGATGAAAA GGCCTGCAGT GATGAGTTGA CTGTGTACAA AGTACAGAA T	4680
	CTTCAGT GG A CAGCTGACTT CTCTGGGAT GTGACTTTGA CCTGGATGAG GCCCAAAAAA	4740
	ATGCCCTCTG CATCTTGTGT ATATAATGTC TACTACAGGG TGGTGGAGA GAGCATATGG	4800
50	AAGACTCTGG AGACCCACAG CAATAAGACA AACACTGTAT TAAAAGTCTT GAAACCAGAT	4860
	ACCACGTATC AGGTTAAAGT ACAGGTT CAG TGTCTCAGCA AGGCACACAA CACCAATGAC	4920

1	TTTGTGACCC TGAGGACCCC AGAGGGATTG CCAGATGCC CTCGAAATCT CCAGCTGTCA	4980
5	CTCCCCAGGG AAGCAGAAGG TGTGATTGTA GGCCACTGGG CTCCCTCCAT CCACACCCAT	5040
10	GGCCTCATCC GTGAGTACAT TGAGAATAAC AGCAGGAGTG GTCAGAAGAT GTGGGCCTCC	5100
15	CAGAGGGCTG CTAGTAACCT TACAGAAATC AAGAACTTAT TGGTCAACAC TCTATACACC	5160
20	GTCAGAGTGG CTGCGGTGAC TAGTCGTGGA ATAGGAAACT GGAGCGATTC TAAATCCATT	5220
25	ACACCATAA AAGGAAAAGT GATCCCACCA CCAGATATCC ACATTGACAG CTATGGTGA	5280
30	AATTATCTAA GCTTCACCCCT GACCATGGAG AGTGTATCA AGGTGAATGG CTATGGTG	5340
35	AACCTTTCT GGGCATTGAA CACCCACAAG CAAGAGAGGA GAACTTTGAA CTTCCGAGGA	5400
40	AGCATATTGT CACACAAAGT TGGCAATCTG ACAGCTCATA CATCCTATGA GATTCTGCC	5460
45	TGGGCCAAGA CTGACTTGGG GGATAGCCCT CTGGCATTG AGCATGTTAT GACCAGAGGG	5520
50	GTTCGCCAC CTGCACCTAG CCTCAAGGCC AAAGCCATCA ACCAGACTGC AGTGGAAATGT	5580
55	ACCTGGACCG GCCCCCGGAA TGTGGTTAT GGTATTTCT ATGCCACGTC CTTCTTGAC	5640
60	CTCTATCGCA ACCCGAAGAG CTTGACTACT TCACTCCACA ACAAGACGGT CATTGTCAGT	5700
65	AAGGATGAGC AGTATTTGTT TCTGGTCCGT GTAGTGGTAC CCTACCAGGG GCCATCCTCT	5760
70	GACTACGTTG TAGTGAAGAT GATCCCAGGAC AGCAGGCTTC CACCCCGTCA CCTGCATGTG	5820
75	GTTCATACGG GCAAAACCTC CGTGGTCATC AAGTGGGAAT CACCGTATGA CTCTCCTGAC	5880
80	CAGGACTTGT TGTATGCAAT TGCAGTCAAA GATCTCATAA GAAAGACTGA CAGGAGCTAC	5940
85	AAAGTAAAAT CCCGTAACAG CACTGTGAA TACACCTTA ACAAGTTGGA GCCTGGCGGG	6000
90	AAATACCACA TCATTGTCCA ACTGGGGAAC ATGAGCAAAG ATTCCAGCAT AAAAATTACC	6060
95	ACAGTTTCAT TATCAGCACC TGATGCCTTA AAAATCATAA CAGAAAATGA TCATGTTCTT	6120
100	CTGTTTGGAA AAAGCCTGGC TTTAAAGGAA AAGCATTGAA ATGAAAGCAG GGGCTATGAG	6180
105	ATACACATGT TTGATAGTGC CATGAATATC ACAGCTTACCC TTGGGAATAC TACTGACAAT	6240
110	TTCTTTAAAA TTTCCAACCT GAAGATGGGT CATAATTACA CGTCACCGT CCAAGCAAGA	6300
115	TGCCTTTTG GCAACCAGAT CTGTGGGGAG CCTGCCATCC TGCTGTACGA TGAGCTGGGG	6360
120	TCTGGTGCAG ATGCATCTGC AACGCAGGCT GCCAGATCTA CGGATGTTGC TGCTGTGGTG	6420
125	GTGCCCATCT TATTCTGTAT ACTGCTGAGC CTGGGGGTGG GGTTTGCCAT CCTGTACACG	6480
130	AAGCACCGGA GGCTGCAGAG CAGCTTCACC GCCTTCGCCA ACAGCCACTA CAGCTCCAGG	6540
135	CTGGGGTCCCG CAATCTCTC CTCTGGGGAT GACCTGGGGG AAGATGATGA AGATGCCCT	6600
140	ATGATAACTG GATTTTCAGA TGACGTCCCC ATGGTGATAG CC	6642

(2) INFORMATION FOR SEQ ID NO: 6:

50 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 2214 amino acids
 (B) TYPE: amino acid

(C) STRANDEDNESS:
 (D) TOPOLOGY: linear

5 (ii) MOLECULE TYPE: Protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

10 Met Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe
 1 5 10 15

Thr Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr
 20 25 30

15 Gln Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe
 35 40 45

Leu Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly
 50 55 60

20 Asp Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys
 65 70 75 80

Arg Ser Ala Ala Leu Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val
 85 90 95

25 Ser Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu
 100 105 110

Lys Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala
 115 120 125

30 Arg Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser
 130 135 140

Phe Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser
 145 150 155 160

35 Glu Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg
 165 170 175

Tyr Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp
 180 185 190

40 Phe Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp
 195 200 205

Leu Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg
 210 215 220

45 Ser His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr
 225 230 235 240

Trp Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp
 245 250 255

50 Pro Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser
 260 265 270

Gly Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu
 275 280 285

Asn Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Cln Leu Arg Asp
 290 295 300
 5 Lys Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln
 305 310 315 320
 Gln Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg
 325 330 335
 Ala Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala
 10 340 345 350
 Asp Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn
 355 360 365
 15 Arg Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu
 370 375 380
 Ser Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp
 385 390 395 400
 20 Thr Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg
 405 410 415
 Val Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser
 420 425 430
 25 Met Asn Glu Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly
 435 440 445
 Thr Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys
 450 455 460
 Ile Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg
 465 470 475 480
 30 Leu Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser
 485 490 495
 Lys Glu Ser Ala Pro Gly Leu Ile Ile Ala Thr Gly Ser Val Gly Lys
 500 505 510
 35 Asn Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ser Ala Gly Ala
 515 520 525
 Arg Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Trp Gly Asp
 530 535 540
 40 His Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu
 545 550 555 560
 Leu Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe
 565 570 575
 45 Ser Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu
 580 585 590
 Lys Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His
 595 600 605
 50 Ser Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro
 610 615 620
 Cys Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly

	625	630	635	640
5	Asn Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro 645 650 655			
	His Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val 660 665 670			
10	Ser Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe 675 680 685			
	Lys Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu 690 695 700			
15	Phe Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser 705 710 715 720			
	Thr Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys 725 730 735			
20	Ser Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys 740 745 750			
	Pro Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser 755 760 765			
	Ile Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu 770 775 780			
25	Thr Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn 785 790 795 800			
	Cys Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys 805 810 815			
30	Leu Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu 820 825 830			
	Thr Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp 835 840 845			
35	Val Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp 850 855 860			
	Phe Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala 865 870 875 880			
40	Leu Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly 885 890 895			
	Asp Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala 900 905 910			
45	Tyr His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val 915 920 925			
	Asp Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu 930 935 940			
50	Arg Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu 945 950 955 960			
	Pro His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp 965 970 975			

Asp Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser
 980 985
 5 Gln Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys
 995 1000 1005
 Ile Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg
 1010 1015 1020
 10 Pro Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys
 1025 1030 1035 1040
 Arg Cys Pro Glu Asp Val Ser Ser Val Leu Pro Ser Gly Asp Leu
 1045 1050 1055
 15 Met Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val
 1060 1065 1070
 Lys Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly
 1075 1080 1085
 20 Asn Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly
 1090 1095 1100
 Asp Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp
 1105 1110 1115 1120
 Thr Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr
 1125 1130 1135
 25 Lys Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His
 1140 1145 1150
 Cys Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly
 1155 1160 1165
 30 Met Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg
 1170 1175 1180
 Asp Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu
 1185 1190 1195 1200
 35 Ala Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp
 1205 1210 1215
 Ala Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro
 1220 1225 1230
 40 Val Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr
 1235 1240 1245
 Cys Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp
 1250 1255 1260
 45 Gly Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp
 1265 1270 1275 1280
 Phe Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys
 1285 1290 1295
 50 Asp Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe
 1300 1305 1310

Ala Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Cys Phe
 1315 1320 1325
 5 Gly Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys
 1330 1335 1340
 Asp Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu
 1345 1350 1355 1360
 10 Asn Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys
 1365 1370 1375
 Glu Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn
 1380 1385 1390
 15 Asp Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile
 1395 1400 1405
 Leu Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr
 1410 1415 1420
 20 Arg Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly
 1425 1430 1435 1440
 Tyr Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu
 1445 1450 1455
 25 Ala Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp
 1460 1465 1470
 Arg Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp
 1475 1480 1485
 30 Lys Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala
 1490 1495 1500
 Asn Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln
 1505 1510 1515 1520
 35 Cys Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly
 1525 1530 1535
 Phe Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu
 1540 1545 1550
 40 Leu Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser
 1555 1560 1565
 Gly Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala
 1570 1575 1580
 45 Ser Cys Val Tyr Asn Val Tyr Arg Val Val Gly Glu Ser Ile Trp
 1585 1590 1595 1600
 Lys Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val
 1605 1610 1615
 50 Leu Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu
 1620 1625 1630
 Ser Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu
 1635 1640 1645
 Gly Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu

	1650	1655	1660
	Ala Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His		
5	1665 1670 1675 1680		
	Gly Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys		
	1685 1690 1695		
	Met Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn		
10	1700 1705 1710		
	Leu Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser		
	1715 1720 1725		
	Arg Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys		
	1730 1735 1740		
15	Gly Lys Val Ile Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu		
	1745 1750 1755 1760		
	Asn Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn		
	1765 1770 1775		
20	Gly Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu		
	1780 1785 1790		
	Arg Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly		
	1795 1800 1805		
25	Asn Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr		
	1810 1815 1820		
	Asp Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly		
	1825 1830 1835 1840		
30	Val Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr		
	1845 1850 1855		
	Ala Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile		
	1860 1865 1870		
35	Phe Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu		
	1875 1880 1885		
	Thr Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln		
	1890 1895 1900		
40	Tyr Leu Phe Leu Val Arg Val Val Val Pro Tyr Gln Gly Pro Ser Ser		
	1905 1910 1915 1920		
	Asp Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg		
	1925 1930 1935		
45	His Leu His Val Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp		
	1940 1945 1950		
	Glu Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala		
	1955 1960 1965		
50	Val Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser		
	1970 1975 1980		
	Arg Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly		
	1985 1990 1995 2000		

Lys Tyr His Ile Ile Val Gin Leu Gly Asn Met Ser Lys Asp Ser Ser
 2005 2010 2015
 5 Ile Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile
 2020 2025 2030
 Ile Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu
 2035 2040 2045
 10 Lys Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe
 2050 2055 2060
 Asp Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn
 2065 2070 2075 2080
 15 Phe Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr
 2085 2090 2095
 Val Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala
 2100 2105 2110
 20 Ile Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr
 2115 2120 2125
 Gln Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Val Pro Ile Leu
 2130 2135 2140
 25 Phe Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr
 2145 2150 2155 2160
 Lys His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His
 2165 2170 2175
 30 Tyr Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu
 2180 2185 2190
 Gly Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp
 2195 2200 2205
 35 Val Pro Met Val Ile Ala
 2210

(2) INFORMATION FOR SEQ ID NO: 7:

40 (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 6843 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

45 (ii) MOLECULE TYPE: cDNA to mRNA

50 (ix) FEATURE:
 (A) NAME/KEY: sig peptide
 (B) LOCATION: 81..164
 (C) IDENTIFICATION METHOD: S

55 (ix) FEATURE:
 (A) NAME/KEY: mat peptide
 (B) LOCATION: 165..6722
 (C) IDENTIFICATION METHOD: S

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

5	CCG GCCCAGCGGC TCTCCTGGCC TCGCGCTGCA CATTCTCTCC TGGCGGCGGC GCCACCTGCA GTAGCGTTCG CCCGAACATG Met	23 83 1
10	GCG ACA CGG AGC AGC AGG AGG GAG TCG CGA CTC CCG TTC CTA TTC ACC Ala Thr Arg Ser Ser Arg Arg Glu Ser Arg Leu Pro Phe Leu Phe Thr 5 10 15 CTG GTC GCA CTG CTG CCG CCC GGA GCT CTC TGC GAA GTC TGG ACG CAG Leu Val Ala Leu Leu Pro Pro Gly Ala Leu Cys Glu Val Trp Thr Gln 10 20 25 30 AGG CTG CAC GGC GGC AGC GCG CCC TTG CCC CAG GAC CGG GGC TTC CTC Arg Leu His Gly Gly Ser Ala Pro Leu Pro Gln Asp Arg Gly Phe Leu 35 40 45 GTG GTG CAG GGC GAC CCG GAG CTG CGG CTG TGG GCG CGC GGG GAT Val Val Gln Gly Asp Pro Arg Glu Leu Arg Leu Trp Ala Arg Gly Asp 50 55 60 65 GCC AGG GGG GCG AGC CGC GCG GAC GAG AAG CCG CTC CGG AGG AAA CGG Ala Arg Gly Ala Ser Arg Ala Asp Glu Lys Pro Leu Arg Arg Lys Arg 70 75 80 AGC GCT GCC CTG CAG CCC GAG CCC ATC AAG GTG TAC GGA CAG GTT AGT Ser Ala Ala Gln Pro Glu Pro Ile Lys Val Tyr Gly Gln Val Ser 80 85 90 95 CTG AAT GAT TCC CAC AAT CAG ATG GTG GTG CAC TGG GCT GGA GAG AAA Leu Asn Asp Ser His Asn Gln Met Val Val His Trp Ala Gly Glu Lys 100 105 110 AGC AAC GTG ATC GTG GCC TTG CGA GAT AGC CTG GCA TTG GCG AGG Ser Asn Val Ile Val Ala Leu Ala Arg Asp Ser Leu Ala Leu Ala Arg 115 120 125 CCC AAG AGC AGT GAT GTG TAC GTG TCT TAC GAC TAT GGA AAA TCA TTC Pro Lys Ser Ser Asp Val Tyr Val Ser Tyr Asp Tyr Gly Lys Ser Phe 130 135 140 145 AAG AAA ATT TCA GAC AAG TTA AAC TTT GGC TTG GGA AAT AGG AGT GAA Lys Lys Ile Ser Asp Lys Leu Asn Phe Gly Leu Gly Asn Arg Ser Glu 150 155 160 GCT GTT ATC GCC CAG TTC TAC CAC AGC CCT GCG GAC AAC AAG CGG TAC Ala Val Ile Ala Gln Phe Tyr His Ser Pro Ala Asp Asn Lys Arg Tyr 165 170 175 ATC TTT GCA GAC GCT TAT GCC CAG TAC CTC TGG ATC ACG TTT GAC TTC Ile Phe Ala Asp Ala Tyr Ala Gln Tyr Leu Trp Ile Thr Phe Asp Phe 180 185 190 TGC AAC ACT CTT CAA GGC TTT TCC ATC CCA TTT CGG GCA GCT GAT CTC Cys Asn Thr Leu Gln Gly Phe Ser Ile Pro Phe Arg Ala Ala Asp Leu 195 200 205 CTC CTA CAC AGT AAG GCC TCC AAC CTT CTC TTG GGC TTT GAC AGG TCC Leu Leu His Ser Lys Ala Ser Asn Leu Leu Leu Gly Phe Asp Arg Ser 210 215 220 225 CAC CCC AAC AAG CAG CTG TGG AAG TCA GAT GAC TTT GGC CAG ACC TGG His Pro Asn Lys Gln Leu Trp Lys Ser Asp Asp Phe Gly Gln Thr Tyr 230 235 240 ATC ATG ATT CAG GAA CAT GTC AAG TCC TTT TCT TGG GGA ATT GAT CCC Ile Met Ile Gln Glu His Val Lys Ser Phe Ser Trp Gly Ile Asp Pro 245 250 255 TAT GAC AAA CCA AAT ACC ATC TAC ATT GAA CGA CAC GAA CCC TCT GGC Tyr Asp Lys Pro Asn Thr Ile Tyr Ile Glu Arg His Glu Pro Ser Gly 260 265 270 TAC TCC ACT GTC TTC CGA AGT ACA GAT TTC TTC CAG TCC CGG GAA AAC Tyr Ser Thr Val Phe Arg Ser Thr Asp Phe Phe Gln Ser Arg Glu Asn 275 280 285 50 CAG GAA GTG ATC CTT GAG GAA GTG AGA GAT TTT CAG CTT CGG GAC AAG Gln Glu Val Ile Leu Glu Glu Val Arg Asp Phe Gln Leu Arg Asp Lys 290 295 300 305 55	131 179 227 275 323 371 419 467 515 563 611 659 707 755 803 899 947 995

	TAC ATG TTT GCT ACA AAG GTG GTG CAT CTC TTG GGC ACT GAA CAC CAG Tyr Met Phe Ala Thr Lys Val Val His Leu Leu Gly Ser Glu Gln Gln 310 315 320	1043
5	TCT TCT GTC CAG CTC TGG GTC TCC TTT GGC CGG AAG CCC ATG AGA GCA Ser Ser Val Gln Leu Trp Val Ser Phe Gly Arg Lys Pro Met Arg Ala 325 330 335	1091
	GCC CAG TTT GTC ACA AGA CAT CCT ATT AAT GAA TAT TAC ATC GCA GAT Ala Gln Phe Val Thr Arg His Pro Ile Asn Glu Tyr Tyr Ile Ala Asp 340 345 350	1139
10	GCC TCC GAG GAC CAG GTG TTT GTG TGT GTC AGC CAC AGT AAC AAC CGC Ala Ser Glu Asp Gln Val Phe Val Cys Val Ser His Ser Asn Asn Arg 355 360 365	1187
	ACC AAT TTA TAC ATC TCA GAG GCA GAG GGG CTG AAG TTC TCC CTG TCC Thr Asn Leu Tyr Ile Ser Glu Ala Glu Gly Leu Lys Phe Ser Leu Ser 370 375 380 385	1235
15	TTG GAG AAC GTG CTC TAT TAC AGC CCA GGA GGG GCC GGC AGT GAC ACC Leu Glu Asn Val Leu Tyr Tyr Ser Pro Gly Gly Ala Gly Ser Asp Thr 390 395 400	1283
	TTG GTG AGG TAT TTT GCA AAT GAA CCA TTT GCT GAC TTC CAC CGA GTG Leu Val Arg Tyr Phe Ala Asn Glu Pro Phe Ala Asp Phe His Arg Val 405 410 415	1331
20	GAA GGA TTG CAA GGA GTC TAC ATT GCT ACT CTG ATT AAT GGT TCT ATG Glu Gly Leu Gln Gly Val Tyr Ile Ala Thr Leu Ile Asn Gly Ser Met 420 425 430	1379
	AAT GAG GAG AAC ATG AGA TCG GTC ATC ACC TTT GAC AAA GGG GGA ACC Asn Glu Asn Met Arg Ser Val Ile Thr Phe Asp Lys Gly Gly Thr 435 440 445	1427
25	TGG GAG TTT CTT CAG GCT CCA GCC TTC ACG GGA TAT GGA GAG AAA ATC Trp Glu Phe Leu Gln Ala Pro Ala Phe Thr Gly Tyr Gly Glu Lys Ile 450 455 460 465	1475
	AAT TGT GAG CTT TCC CAG GGC TGT TCC CTT CAT CTG GCT CAG CGC CTC Asn Cys Glu Leu Ser Gln Gly Cys Ser Leu His Leu Ala Gln Arg Leu 470 475 480	1523
30	AGT CAG CTC CTC AAC CTC CAG CTC CGG AGA ATG CCC ATC CTG TCC AAG Ser Gln Leu Leu Asn Leu Gln Leu Arg Arg Met Pro Ile Leu Ser Lys 485 490 495	1571
	GAG TCG GCT CCA GGC CTC ATC ATC GCC ACT GGC TCA GTG GGA AAG AAC Glu Ser Ala Pro Gly Leu Ile Ala Thr Gly Ser Val Gly Lys Asn 500 505 510	1619
35	TTG GCT AGC AAG ACA AAC GTG TAC ATC TCT AGC AGT GCT GGA GCC AGG Leu Ala Ser Lys Thr Asn Val Tyr Ile Ser Ser Ala Gly Ala Arg 515 520 525	1667
	TGG CGA GAG GCA CTT CCT GGA CCT CAC TAC TAC ACA TGG GGA GAC CAC Trp Arg Glu Ala Leu Pro Gly Pro His Tyr Tyr Thr Trp Gly Asp His 530 535 540 545	1715
40	GGC GGA ATC ATC ACG GCC ATT GCC CAG GGC ATG GAA ACC AAC GAG CTA Gly Gly Ile Ile Thr Ala Ile Ala Gln Gly Met Glu Thr Asn Glu Leu 550 555 560	1763
	AAA TAC AGT ACC AAT GAA GGG GAG ACC TGG AAA ACA TTC ATC TTC TCT Lys Tyr Ser Thr Asn Glu Gly Glu Thr Trp Lys Thr Phe Ile Phe Ser 565 570 575	1811
45	GAG AAG CCA GTG TTT GTG TAT GGC CTC CTC ACA GAA CCT GGG GAG AAG Glu Lys Pro Val Phe Val Tyr Gly Leu Leu Thr Glu Pro Gly Glu Lys 580 585 590	1859
	AGC ACT GTC TTC ACC ATC TTT GGC TCG AAC AAA GAG AAT GTC CAC AGC Ser Thr Val Phe Thr Ile Phe Gly Ser Asn Lys Glu Asn Val His Ser 595 600 605	1907
50	TGG CTG ATC CTC CAG GTC AAT GCC ACG GAT GCC TTG GGA GTT CCC TGC Trp Leu Ile Leu Gln Val Asn Ala Thr Asp Ala Leu Gly Val Pro Cys 610 615 620 625	1955
	ACA GAG AAT GAC TAC AAG CTC TGG TCA CCA TCT GAT GAG CGG GGG AAT Thr Glu Asn Asp Tyr Lys Leu Trp Ser Pro Ser Asp Glu Arg Gly Asn 630 635 640	2003
	GAG TGT TTG CTG GGA CAC AAG ACT GTT TTC AAA CGG CGG ACC CCC CAT	2051

	Glu Cys Leu Leu Gly His Lys Thr Val Phe Lys Arg Arg Thr Pro His		
	645 650 655		
5	CCC ACA TGC TTC AAT GGA GAG GAC TTT GAC AGG CCG GTG GTC GTG TCC	2099	
	Ala Thr Cys Phe Asn Gly Glu Asp Phe Asp Arg Pro Val Val Val Ser		
	660 665 670		
	AAC TGC TCC TGC ACC CGG GAG GAC TAT GAG TGT GAC TTC GGT TTC AAG	2147	
	Asn Cys Ser Cys Thr Arg Glu Asp Tyr Glu Cys Asp Phe Gly Phe Lys		
	675 680 685		
10	ATG AGT GAA GAT TTG TCA TTA GAG GTT TGT GTT CCA GAT CCG GAA TTT	2195	
	Met Ser Glu Asp Leu Ser Leu Glu Val Cys Val Pro Asp Pro Glu Phe		
	690 695 700 705		
	TCT GGA AAG TCA TAC TCC CCT CCT GTG CCT TGC CCT GTG GGT TCT ACT	2243	
	Ser Gly Lys Ser Tyr Ser Pro Pro Val Pro Cys Pro Val Gly Ser Thr		
	710 715 720		
15	TAC AGG AGA ACG AGA GGC TAC CGG AAG ATT TCT GGG GAC ACT TGT AGC	2291	
	Tyr Arg Arg Thr Arg Gly Tyr Arg Lys Ile Ser Gly Asp Thr Cys Ser		
	725 730 735		
	GGA GGA GAT GTT GAA GCG CGA CTG GAA GGA GAG CTG GTC CCC TGT CCC	2339	
	Gly Gly Asp Val Glu Ala Arg Leu Glu Gly Glu Leu Val Pro Cys Pro		
	740 745 750		
20	CTG GCA GAA GAG AAC GAG TTC ATT CTG TAT GCT GTG AGG AAA TCC ATC	2387	
	Leu Ala Glu Glu Asn Glu Phe Ile Leu Tyr Ala Val Arg Lys Ser Ile		
	755 760 765		
	TAC CGC TAT GAC CTG GCC TCG GGA GCC ACC GAG CAG TTG CCT CTC ACC	2435	
	Tyr Arg Tyr Asp Leu Ala Ser Gly Ala Thr Glu Gln Leu Pro Leu Thr		
	770 775 780 785		
	GGG CTA CGG GCA GCA GTG GCC CTG GAC TTT GAC TAT GAG CAC AAC TGT	2483	
	Gly Leu Arg Ala Ala Val Ala Leu Asp Phe Asp Tyr Glu His Asn Cys		
	790 795 800		
25	TTG TAT TGG TCC GAC CTG GCC TTG GAC GTC ATC CAG CGC CTC TGT TTG	2531	
	Leu Tyr Trp Ser Asp Leu Ala Leu Asp Val Ile Gln Arg Leu Cys Leu		
	805 810 815		
	AAT GGA AGC ACA GGG CAA GAG GTG ATC ATC AAT TCT GGC CTG GAG ACA	2579	
	Asn Gly Ser Thr Gly Gln Glu Val Ile Ile Asn Ser Gly Leu Glu Thr		
	820 825 830		
30	GTA GAA GCT TTG GCT TTT GAA CCC CTC AGC CAG CTG CTT TAC TGG GTA	2627	
	Val Glu Ala Leu Ala Phe Glu Pro Leu Ser Gln Leu Leu Tyr Trp Val		
	835 840 845		
	GAT GCA GGC TTC AAA AAG ATT GAG GTA GCT AAT CCA GAT GGC GAC TTC	2675	
	Asp Ala Gly Phe Lys Lys Ile Glu Val Ala Asn Pro Asp Gly Asp Phe		
	850 855 860 865		
35	CGA CTC ACA ATC GTC AAT TCC TCT GTG CTT GAT CGT CCC AGG GCT CTG	2723	
	Arg Leu Thr Ile Val Asn Ser Ser Val Leu Asp Arg Pro Arg Ala Leu		
	870 875 880		
	GTC CTC GTG CCC CAA GAG GGG GTG ATG TTC TGG ACA GAC TGG GGA GAC	2771	
	Val Leu Val Pro Gln Glu Gly Val Met Phe Trp Thr Asp Trp Gly Asp		
	885 890 895		
40	CTG AAG CCT GGG ATT TAT CGG AGC AAT ATG GAT GGT TCT GCT GCC TAT	2819	
	Leu Lys Pro Gly Ile Tyr Arg Ser Asn Met Asp Gly Ser Ala Ala Tyr		
	900 905 910		
	CAC CTG GTG TCT GAG GAT GTG AAG TGG CCC AAT GGC ATC TCT GTG GAC	2867	
	His Leu Val Ser Glu Asp Val Lys Trp Pro Asn Gly Ile Ser Val Asp		
	915 920 925		
45	GAC CAG TGG ATT TAC TGG ACG GAT GCC TAC CTG GAG TGC ATA GAG CGG	2915	
	Asp Gln Trp Ile Tyr Trp Thr Asp Ala Tyr Leu Glu Cys Ile Glu Arg		
	930 935 940 945		
	ATC ACG TTC AGT GGC CAG CAG CGC TCT GTC ATT CTG GAC AAC CTC CCG	2963	
	Ile Thr Phe Ser Gly Gln Gln Arg Ser Val Ile Leu Asp Asn Leu Pro		
	950 955 960		
50	CAC CCC TAT GCC ATT GCT GTC TTT AAG AAT GAA ATC TAC TGG GAT GAC	3011	
	His Pro Tyr Ala Ile Ala Val Phe Lys Asn Glu Ile Tyr Trp Asp Asp		
	965 970 975		
	TGG TCA CAG CTC AGC ATA TTC CGA GCT TCC AAA TAC AGT GGG TCC CAG	3059	
	Trp Ser Gln Leu Ser Ile Phe Arg Ala Ser Lys Tyr Ser Gly Ser Gln		

	980	985	990	
5	ATG GAG ATT CTG GCA AAC CAG CTC ACG GG; CTC AT; GAC ATG AAG AAT Met Glu Ile Leu Ala Asn Gln Leu Thr Gly Leu Met Asp Met Lys Ile 995 1000 1005			3107
	TTC TAC AAG GGG AAG AAC ACT GGA AGC AAT GCC TGT GTG CCC AGG CCA Phe Tyr Lys Gly Lys Asn Thr Gly Ser Asn Ala Cys Val Pro Arg Pro 1010 1015 1020 1025			3155
	TGC AGC CTG CTG TGC CCC AAG GCC AAC AAC AGT AGA AGC TGC AGG Cys Ser Leu Leu Cys Leu Pro Lys Ala Asn Asn Ser Arg Ser Cys Arg 1030 1035 1040			3203
10	TGT CCA GAG GAT GTG TCC AGC AGT GTG CTT CCA TCA GGG GAC CTG ATG Cys Pro Glu Asp Val Ser Ser Val Leu Pro Ser Gly Asp Leu Met 1045 1050 1055			3251
	TGT GAC TGC CCT CAG GGC TAT CAG CTC AAG AAC AAT ACC TGT GTC AAA Cys Asp Cys Pro Gln Gly Tyr Gln Leu Lys Asn Asn Thr Cys Val Lys 1060 1065 1070			3299
15	GAA GAG AAC ACC TGT CTT CGC AAC CAG TAT CGC TGC AGC AAC GGG AAC Glu Glu Asn Thr Cys Leu Arg Asn Gln Tyr Arg Cys Ser Asn Gly Asn 1075 1080 1085			3347
	TGT ATC AAC AGC ATT TGG TGG TGT GAC TTT GAC AAC GAC TGT GGA GAC Cys Ile Asn Ser Ile Trp Trp Cys Asp Phe Asp Asn Asp Cys Gly Asp 1090 1095 1100 1105			3395
20	ATG AGC GAT GAG AGA AAC TGC CCT ACC ACC ATC TGT GAC CTG GAC ACC Met Ser Asp Glu Arg Asn Cys Pro Thr Thr Ile Cys Asp Leu Asp Thr 1110 1115 1120			3443
	CAG TTT CGT TGC CAG GAG TCT GGG ACT TGT ATC CCA CTG TCC TAT AAA Gln Phe Arg Cys Gln Glu Ser Gly Thr Cys Ile Pro Leu Ser Tyr Lys 1125 1130 1135			3491
25	TGT GAC CTT GAG GAT GAC TGT GGA GAC AAC AGT GAT GAA AGT CAT TGT Cys Asp Leu Glu Asp Asp Cys Gly Asp Asn Ser Asp Glu Ser His Cys 1140 1145 1150			3539
	GAA ATG CAC CAG TGC CGG AGT GAC GAG TAC AAC TGC AGT TCC GGC ATG Glu Met His Gln Cys Arg Ser Asp Glu Tyr Asn Cys Ser Ser Gly Met 1155 1160 1165			3587
30	TGC ATC CGC TCC TCC TGG GTA TGT GAC GGG GAC AAC GAC TGC AGG GAC Cys Ile Arg Ser Ser Trp Val Cys Asp Gly Asp Asn Asp Cys Arg Asp 1170 1175 1180 1185			3635
	TGG TCT GAT GAA GCC AAC TGT ACC GCC ATC TAT CAC ACC TGT GAG GCC Trp Ser Asp Glu Ala Asn Cys Thr Ala Ile Tyr His Thr Cys Glu Ala 1190 1195 1200			3683
35	TCC AAC TTC CAG TGC CGA AAC GGG CAC TGC ATC CCC CAG CGG TGG GCG Ser Asn Phe Gln Cys Arg Asn Gly His Cys Ile Pro Gln Arg Trp Ala 1205 1210 1215			3731
	TGT GAC GGG GAT ACG GAC TGC CAG GAT GGT TCC GAT GAG GAT CCA GTC Cys Asp Gly Asp Thr Asp Cys Gln Asp Gly Ser Asp Glu Asp Pro Val 1220 1225 1230 1235			3779
40	AAC TGT GAG AAG AAG TGC AAT GGA TTC CGC TGC CCA AAC GGC ACT TGC Asn Cys Glu Lys Lys Cys Asn Gly Phe Arg Cys Pro Asn Gly Thr Cys 1235 1240 1245			3827
	ATC CCA TCC AGC AAA CAT TGT GAT GGT CTG CGT GAT TGC TCT GAT GGC Ile Pro Ser Ser Lys His Cys Asp Gly Leu Arg Asp Cys Ser Asp Gly 1250 1255 1260 1265			3875
45	TCC GAT GAA CAG CAC TGC GAG CCC CTC TGT ACG CAC TTC ATG GAC TTT Ser Asp Glu Gln His Cys Glu Pro Leu Cys Thr His Phe Met Asp Phe 1270 1275 1280			3923
	GTG TGT AAG AAC CGC CAG CAG TGC CTG TTC CAC TCC ATG GTC TGT GAC Val Cys Lys Asn Arg Gln Gln Cys Leu Phe His Ser Met Val Cys Asp 1285 1290 1295			3971
50	GGA ATC ATC CAG TGC CGC GAC GGG TCC GAT GAG GAT GCG GCG TTT GCA Gly Ile Ile Gln Cys Arg Asp Gly Ser Asp Glu Asp Ala Ala Phe Ala 1300 1305 1310			4019
	GGA TGC TCC CAA GAT CCT GAG TTC CAC AAG GTA TGT GAT GAG TTC GGT Gly Cys Ser Gln Asp Pro Glu Phe His Lys Val Cys Asp Glu Phe Gly 1315 1320 1325			4067

	TTC CAG TGT CAG AAT GGA GTG TGC ATC AGT TTG ATT TGG AAG TGC GAC	4115
	Phe Gln Cys Gln Asn Gly Val Cys Ile Ser Leu Ile Trp Lys Cys Asp	
5	1330 1335 1340 1345	
	GGG ATG GAT GAT TGC GGC GAT 'AT TCT GAT GAA GCC AAC TGC GAA AAC	4163
	Gly Met Asp Asp Cys Gly Asp Tyr Ser Asp Glu Ala Asn Cys Glu Asn	
	1350 1355 1360	
	CCC ACA GAA GCC CCA AAC TGC TCC CGC TAC TTC CAG TTT CGG TGT GAG	4211
	Pro Thr Glu Ala Pro Asn Cys Ser Arg Tyr Phe Gln Phe Arg Cys Glu	
	1365 1370 1375	
10	AAT GGC CAC TGC ATC CCC AAC AGA TGG AAA TGT GAC AGG GAG AAC GAC	4259
	Asn Gly His Cys Ile Pro Asn Arg Trp Lys Cys Asp Arg Glu Asn Asp	
	1380 1385 1390	
	TGT GGG GAC TGG TCT GAT GAG AAG GAT TGT GGA GAT TCA CAT ATT CTT	4307
	Cys Gly Asp Trp Ser Asp Glu Lys Asp Cys Gly Asp Ser His Ile Leu	
	1395 1400 1405	
15	CCC TTC TCG ACT CCT GGG CCC TCC ACG TGT CTG CCC AAT TAC TAC CGC	4355
	Pro Phe Ser Thr Pro Gly Pro Ser Thr Cys Leu Pro Asn Tyr Tyr Arg	
	1410 1415 1420 1425	
	TGC AGC AGT GGG ACC TGC GTG ATG GAC ACC TGG GTG TGC GAC GGG TAC	4403
	Cys Ser Ser Gly Thr Cys Val Met Asp Thr Trp Val Cys Asp Gly Tyr	
	1430 1435 1440	
20	CGA GAT TGT GCA GAT GGC TCT GAC GAG GAA GCC TGC CCC TTG CTT GCA	4451
	Arg Asp Cys Ala Asp Gly Ser Asp Glu Glu Ala Cys Pro Leu Leu Ala	
	1445 1450 1455	
	AAC GTC ACT GCT GCC TCC ACT CCC ACC CAA CTT GGG CGA TGT GAC CGA	4499
	Asn Val Thr Ala Ala Ser Thr Pro Thr Gln Leu Gly Arg Cys Asp Arg	
	1460 1465 1470	
25	TTT GAG TTC GAA TGC CAC CAA CCG AAG ACG TGT ATT CCC AAC TGG AAG	4547
	Phe Glu Phe Glu Cys His Gln Pro Lys Thr Cys Ile Pro Asn Trp Lys	
	1475 1480 1485	
	CGC TGT GAC GGC CAC CAA GAT TGC CAG GAT GGC CGG GAC GAG GCC AAT	4595
	Arg Cys Asp Gly His Gln Asp Cys Gln Asp Gly Arg Asp Glu Ala Asn	
	1490 1495 1500 1505	
	TGC CCC ACA CAC AGC ACC TTG ACT TGC ATG AGC AGG GAG TTC CAG TGC	4643
	Cys Pro Thr His Ser Thr Leu Thr Cys Met Ser Arg Glu Phe Gln Cys	
30	1510 1515 1520	
	GAG GAC GGG GAG GCC TGC ATT GTG CTC TCG GAG CGC TGC GAC GGC TTC	4691
	Glu Asp Gly Glu Ala Cys Ile Val Leu Ser Glu Arg Cys Asp Gly Phe	
	1525 1530 1535	
	CTG GAC TGC TCG GAC GAG AGC GAT GAA AAG GCC TGC AGT GAT GAG TTG	4739
	Leu Asp Cys Ser Asp Glu Ser Asp Glu Lys Ala Cys Ser Asp Glu Leu	
35	1540 1545 1550	
	ACT GTG TAC AAA GTA CAG AAT CTT CAG TGG ACA GCT GAC TTC TCT GGG	4787
	Thr Val Tyr Lys Val Gln Asn Leu Gln Trp Thr Ala Asp Phe Ser Gly	
	1555 1560 1565	
	GAT GTG ACT TTG ACC TGG ATG AGG CCC AAA AAA ATG CCC TCT GCA TCT	4835
	Asp Val Thr Leu Thr Trp Met Arg Pro Lys Lys Met Pro Ser Ala Ser	
40	1570 1575 1580 1585	
	TGT GTA TAT AAT GTC TAC TAC AGG GTG GTT GGA GAG AGC ATA TGG AAG	4883
	Cys Val Tyr Asn Val Tyr Tyr Arg Val Val Gly Glu Ser Ile Trp Lys	
	1590 1595 1600	
	ACT CTG GAG ACC CAC AGC AAT AAG ACA AAC ACT GTA TTA AAA GTC TTG	4931
	Thr Leu Glu Thr His Ser Asn Lys Thr Asn Thr Val Leu Lys Val Leu	
	1605 1610 1615	
45	AAA CCA GAT ACC ACG TAT CAG GTT AAA GTA CAG GTT CAG TGT CTC AGC	4979
	Lys Pro Asp Thr Thr Tyr Gln Val Lys Val Gln Val Gln Cys Leu Ser	
	1620 1625 1630	
	AAG GCA CAC ACC AAT GAC TTT GTG ACC CTG AGG ACC CCA GAG GGA	5027
	Lys Ala His Asn Thr Asn Asp Phe Val Thr Leu Arg Thr Pro Glu Gly	
	1635 1640 1645	
50	TTG CCA GAT GCC CCT CGA AAT CTC CAG CTG TCA CTC CCC AGG GAA GCA	5075
	Leu Pro Asp Ala Pro Arg Asn Leu Gln Leu Ser Leu Pro Arg Glu Ala	
	1650 1655 1660 1665	
	GAA GGT GTG ATT GTA GGC CAC TGG GCT CCT CCC ATC CAC ACC CAT GGC	5123

	Glu Gly Val Ile Val Gly His Trp Ala Pro Pro Ile His Thr His Gly		
	1670 1675 1680		
5	CTC ATC CGT GAG TAC ATT GTA GAA TAC AGC AGG AGT GGT TCC AAG ATG	5171	
	Leu Ile Arg Glu Tyr Ile Val Glu Tyr Ser Arg Ser Gly Ser Lys Met		
	1685 1690 1695		
	TGG GCC TCC CAG AGG GCT GCT AGT AAC TTT ACA GAA ATC AAG AAC TTA	5219	
	Trp Ala Ser Gln Arg Ala Ala Ser Asn Phe Thr Glu Ile Lys Asn Leu		
	1700 1705 1710		
10	TTG GTC AAC ACT CTA TAC ACC GTC AGA GTG GCT GCG GTG ACT AGT CGT	5267	
	Leu Val Asn Thr Leu Tyr Thr Val Arg Val Ala Ala Val Thr Ser Arg		
	1715 1720 1725		
	GGA ATA GGA AAC TGG AGC GAT TCT AAA TCC ATT ACC ACC ATA AAA GGA	5315	
	Gly Ile Gly Asn Trp Ser Asp Ser Lys Ser Ile Thr Thr Ile Lys Gly		
	1730 1735 1740 1745		
15	AAA GTG ATC CCA CCA CCA GAT ATC CAC ATT GAC AGC TAT GGT GAA AAT	5363	
	Lys Val Ile Pro Pro Asp Ile His Ile Asp Ser Tyr Gly Glu Asn		
	1750 1755 1760		
	TAT CTA AGC TTC ACC CTG ACC ATG GAG AGT GAT ATC AAG GTG AAT GGC	5411	
	Tyr Leu Ser Phe Thr Leu Thr Met Glu Ser Asp Ile Lys Val Asn Gly		
	1765 1770 1775		
20	TAT GTG GTG AAC CTT TTC TGG GCA TTT GAC ACC CAC AAG CAA GAG AGG	5459	
	Tyr Val Val Asn Leu Phe Trp Ala Phe Asp Thr His Lys Gln Glu Arg		
	1780 1785 1790		
	AGA ACT TTG AAC TTC CGA GGA AGC ATA TTG TCA CAC AAA GTT GGC AAT	5507	
	Arg Thr Leu Asn Phe Arg Gly Ser Ile Leu Ser His Lys Val Gly Asn		
	1795 1800 1805		
25	CTG ACA GCT CAT ACA TCC TAT GAG ATT TCT GCC TGG GCC AAG ACT GAC	5555	
	Leu Thr Ala His Thr Ser Tyr Glu Ile Ser Ala Trp Ala Lys Thr Asp		
	1810 1815 1820 1825		
	TTG GGG GAT AGC CCT CTG GCA TTT GAG CAT GTT ATG ACC AGA GGG GTT	5603	
	Leu Gly Asp Ser Pro Leu Ala Phe Glu His Val Met Thr Arg Gly Val		
	1830 1835 1840		
	CGC CCA CCT GCA CCT AGC CTC AAG GCC AAA GCC ATC AAC CAG ACT GCA	5651	
	Arg Pro Pro Ala Pro Ser Leu Lys Ala Lys Ala Ile Asn Gln Thr Ala		
	1845 1850 1855		
30	GTG GAA TGT ACC TGG ACC GGC CCC CGG AAT GTG GTT TAT GGT ATT TTC	5699	
	Val Glu Cys Thr Trp Thr Gly Pro Arg Asn Val Val Tyr Gly Ile Phe		
	1860 1865 1870		
	TAT GCC ACG TCC TTT CTT GAC CTC TAT CGC AAC CCG AAG AGC TTG ACT	5747	
	Tyr Ala Thr Ser Phe Leu Asp Leu Tyr Arg Asn Pro Lys Ser Leu Thr		
	1875 1880 1885		
35	ACT TCA CTC CAC AAC AAG ACG GTC ATT GTC AGT AAG GAT GAG CAG TAT	5795	
	Thr Ser Leu His Asn Lys Thr Val Ile Val Ser Lys Asp Glu Gln Tyr		
	1890 1895 1900 1905		
	TTG TTT CTG GTC CGT GTA GTG GTA CCC TAC CAG GGG CCA TCC TCT GAC	5843	
	Leu Phe Leu Val Arg Val Val Pro Tyr Gln Gly Pro Ser Ser Asp		
	1910 1915 1920		
40	TAC GTT GTA GTG AAG ATG ATC CCG GAC AGC AGG CTT CCA CCC CGT CAC	5891	
	Tyr Val Val Val Lys Met Ile Pro Asp Ser Arg Leu Pro Pro Arg His		
	1925 1930 1935		
	CTG CAT GTG GTT CAT ACG GGC AAA ACC TCC GTG GTC ATC AAG TGG GAA	5939	
	Leu His Val His Thr Gly Lys Thr Ser Val Val Ile Lys Trp Glu		
	1940 1945 1950		
45	TCA CCG TAT GAC TCT CCT GAC CAG GAC TTG TTG TAT GCA ATT GCA GTC	5987	
	Ser Pro Tyr Asp Ser Pro Asp Gln Asp Leu Leu Tyr Ala Ile Ala Val		
	1955 1960 1965		
	AAA GAT CTC ATA AGA AAG ACT GAC AGG AGC TAC AAA GTA AAA TCC CGT	6035	
	Lys Asp Leu Ile Arg Lys Thr Asp Arg Ser Tyr Lys Val Lys Ser Arg		
	1970 1975 1980 1985		
50	AAC AGC ACT GTG GAA TAC ACC CTT AAC AAG TTG GAG CCT GGC GGG AAA	6083	
	Asn Ser Thr Val Glu Tyr Thr Leu Asn Lys Leu Glu Pro Gly Gly Lys		
	1990 1995 2000		
	TAC CAC ATC ATT GTC CAA CTG GGG AAC ATG AGC AAA GAT TCC AGC ATA	6131	
	Tyr His Ile Ile Val Gln Leu Gly Asn Met Ser Lys Asp Ser Ser Ile		

	2005	2010	2015	
5	AAA ATT ACC ACA GTT TCA TCA GCA CCT' GAT GCG' TTA AAA ATC A' A Lys Ile Thr Thr Val Ser Leu Ser Ala Pro Asp Ala Leu Lys Ile Ile 2020 2025 2030			6179
10	ACA GAA AAT GAT CAT GTT CTT CTG TTT TGG AAA AGC CTG GCT TTA AAG Thr Glu Asn Asp His Val Leu Leu Phe Trp Lys Ser Leu Ala Leu Lys 2035 2040 2045			6227
15	GAA AAG CAT TTT AAT GAA AGC AGG GGC TAT GAG ATA CAC ATG TTT GAT Glu Lys His Phe Asn Glu Ser Arg Gly Tyr Glu Ile His Met Phe Asp 2050 2055 2060 2065			6275
20	AGT GCC ATG AAT ATC ACA GCT TAC CTT GGG AAT ACT ACT GAC AAT TTC Ser Ala Met Asn Ile Thr Ala Tyr Leu Gly Asn Thr Thr Asp Asn Phe 2070 2075 2080			6323
25	TTT AAA ATT TCC AAC CTG AAG ATG GGT CAT AAT TAC ACG TTC ACC GTC Phe Lys Ile Ser Asn Leu Lys Met Gly His Asn Tyr Thr Phe Thr Val 2085 2090 2095			6371
30	CAA GCA AGA TGC CTT TTT GGC AAC CAG ATC TGT GGG GAG CCT GCC ATC Gln Ala Arg Cys Leu Phe Gly Asn Gln Ile Cys Gly Glu Pro Ala Ile 2100 2105 2110			6419
35	CTG CTG TAC GAT GAG CTG GGG TCT GGT GCA GAT GCA TCT GCA ACG CAG Leu Leu Tyr Asp Glu Leu Gly Ser Gly Ala Asp Ala Ser Ala Thr Gln 2115 2120 2125			6467
40	GCT GCC AGA TCT ACG GAT GTT GCT GCT GTG GTG CCC ATC TTA TTC Ala Ala Arg Ser Thr Asp Val Ala Ala Val Val Pro Ile Leu Phe 2130 2135 2140 2145			6515
45	CTG ATA CTG CTG AGC CTG GGG GTG GGG TTT GCC ATC CTG TAC ACG AAG Leu Ile Leu Leu Ser Leu Gly Val Gly Phe Ala Ile Leu Tyr Thr Lys 2150 2155 2160			6563
50	CAC CGG AGG CTG CAG AGC AGC TTC ACC GCC TTC GCC AAC AGC CAC TAC His Arg Arg Leu Gln Ser Ser Phe Thr Ala Phe Ala Asn Ser His Tyr 2165 2170 2175			6611
55	AGC TCC AGG CTG GGG TCC GCA ATC TTC TCC TCT GGG GAT GAC CTG GGG Ser Ser Arg Leu Gly Ser Ala Ile Phe Ser Ser Gly Asp Asp Leu Gly 2180 2185 2190			6659
60	GAA GAT GAT GAA GAT GCC CCT ATG ATA ACT GGA TTT TCA GAT GAC GTC Glu Asp Asp Glu Asp Ala Pro Met Ile Thr Gly Phe Ser Asp Asp Val 2195 2200 2205			6707
65	CCC ATG GTG ATA GCC TGAAAGAGCT TTCCTCACTA GAAACCAAAT GGTGTAAATA Pro Met Val Ile Ala 2210			6762
70	TTTTATTTGA TAAAGATAGT TGATGGTTA TTTTAAAAGA TGCACTTGA GTTGCAATAT GTTATTTTA TATGGGCCAA A			6822 6843

40

Claims

45

1. DNA having a nucleotide sequence as shown by Sequence ID No. 1.
2. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 2 and coded by the DNA of Claim 1.
3. DNA having a nucleotide sequence as shown by Sequence ID No. 5
4. An LDL receptor analog protein having an amino acid sequence as shown by Sequence ID No. 6 and coded by the DNA of Claim 3.
5. A recombinant vector comprising DNA as shown by Sequence ID No. 1 or 5 and a replicable vector.
6. Transformant cells which harbor the recombinant vector of Claim 5.

7. A method for the production of an LDL receptor analog protein comprising the steps of culturing the transformants of Claim 6 and collecting a polypeptide produced in the culture.

14
15

10

15

20

25

30

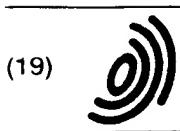
35

40

45

50

55



(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets



(11)

EP 0 773 290 A3

(12)

EUROPEAN PATENT APPLICATION

(88) Date of publication A3:
08.07.1998 Bulletin 1998/28

(51) Int. Cl.⁶: C12N 15/12, C07K 14/705
// C12N15/70, C12N15/79

(43) Date of publication A2:
14.05.1997 Bulletin 1997/20

(21) Application number: 96116108.0

(22) Date of filing: 08.10.1996

(84) Designated Contracting States:
AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC
NL PT SE

• Iwasaki, Akio
Tsuchiura-shi, Ibaraki (JP)
• Arai, Koichi
Urawa-shi, Saitama (JP)
• Yamazaki, Hiroyuki
Higashimurayama-shi, Tokyo (JP)

(30) Priority: 09.10.1995 JP 261440/95
24.04.1996 JP 102451/96

(71) Applicant: KOWA COMPANY, LTD.
Naka-ku Nagoya-shi Aichi-ken (JP)
(74) Representative:
Wächtershäuser, Günter, Prof. Dr.
Patentanwalt,
Tal 29
80331 München (DE)

(72) Inventors:
• Saito, Yasushi
Chiba-shi, Chiba (JP)

(54) Novel LDL receptor analog protein and the gene coding therefor

(57) The present invention is drawn to the gene of a novel LDL receptor family receptor which participates in lipoprotein metabolism, a critical factor that triggers the onset of arteriosclerosis.

The invention provides DNA having a nucleotide sequence as shown by Sequence ID No. 1 or No 5 is disclosed as well as rabbit tissue or human tissue LDL receptor analog protein having an amino acid sequence of Sequence ID No. 2 or 6 coded by such DNA.

EP 0 773 290 A3



DOCUMENTS CONSIDERED TO BE RELEVANT			CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	
A	WO 94 10322 A (UNIVERSITY OF TEXAS) * the whole document * ---	1-7	C12N15/12 C07K14/705 //C12N15/70, C12N15/79
A	WO 95 13374 A (BAYLOR COLLEGE OF MEDICINE) * the whole document * ---	1-7	
A	KOZARSKY K. ET AL.: "In vivo correction of low density lipoprotein receptor deficiency in the Watanabe heritable hyperlipemic rabbit with recombinant adenoviruses" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 269, no. 18, 6 May 1994, MD US, pages 13695-13702, XP002062181 * the whole document * ---	1-7	
A	WILSON J.M. ET AL.: "Hepatocyte-directed gene transfer in vivo leads to transient improvement of hypercholesterolemia in low density lipoprotein receptor-deficient rabbits" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 267, no. 2, 15 January 1995, MD US, pages 963-967, XP002062183 * the whole document * -----	1-7	
The present search report has been drawn up for all claims			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			C07K
Place of search BERLIN			Examiner Panzica, G
CATEGORY OF CITED DOCUMENTS			
X particularly relevant if taken alone Y particularly relevant if combined with another document of the same category A technological background O non-written disclosure P intermediate document			
T theory or principle underlying the invention E earlier patent document, but published on, or after the filing date D document cited in the application L document cited for other reasons & member of the same patent family, corresponding document			